

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:12 : Search time 61.04 Seconds
(without alignments)
331.593 Million cell updates/sec

Title: US-09-119-209-2_COPY_39_155

Perfect score: 666
Sequence: 1 WYHYSEKMNWGRARFCR.....NMDAGKMNDDACKLKALC 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPRMBL_19: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp Unclassified: *
15: sp_virus: *
16: sp_bacteria: *
17: sp_archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	385	4 Q9UJ43	Q9UJ43 homo sapien
2	598	89.8	376	6 Q28629	Q28629 oryctolagus
3	557	83.6	372	11 Q63762	Q63762 rattus norv
4	445	66.8	616	4 Q95509	Q95509 homo sapien
5	445	66.8	740	4 Q95507	Q95507 homo sapien
6	445	66.8	740	4 Q95508	Q95508 homo sapien
7	440	66.1	484	6 Q951G2	Q951G2 ovls aries
8	437	65.6	754	6 Q951G3	Q951G3 odocoileus
9	434	65.2	785	6 Q28290	Q28290 canis faml1
10	425	63.8	482	6 Q28982	Q28982 sus scrofa
11	420	63.1	649	6 Q28657	Q28657 oryctolagus
12	408	61.3	609	6 Q9GLR0	Q9GLR0 canis faml1
13	401	60.2	646	6 Q29097	Q29097 sus scrofa
14	392.5	58.9	610	6 Q951G1	Q951G1 equus caball
15	177	26.6	328	4 Q9UJ71	Q9UJ71 homo sapien
16	166	24.9	359	5 Q9VQX3	Q9VQX3 drosophila

17	162.5	24.4	381	6 Q95LA8	Q95LA8 macaca mula
18	161.5	24.2	268	4 Q96Q03	Q96Q03 homo sapien
19	161.5	24.2	312	4 Q96Q07	Q96Q07 homo sapien
20	161.5	24.2	334	4 Q96Q09	Q96Q09 homo sapien
21	161.5	24.2	360	4 Q96Q04	Q96Q04 homo sapien
22	161.5	24.2	380	4 Q96Q05	Q96Q05 homo sapien
23	161.5	24.2	380	4 Q96Q00	Q96Q00 homo sapien
24	161.5	24.2	381	6 Q95LC6	Q95LC6 macaca neme
25	161.5	24.2	404	4 Q9NNX6	Q9NNX6 homo sapien
26	161.5	24.2	404	4 Q96Q01	Q96Q01 homo sapien
27	160.5	24.1	292	4 Q14538	Q14538 homo sapien
28	160.5	24.1	404	6 Q95L98	Q95L98 pan troglod
29	160.5	24.1	404	6 Q95J96	Q95J96 macaca mula
30	153.5	23.0	304	11 Q91YT3	Q91YT3 mus musculu
31	150.5	22.6	287	4 Q03969	Q03969 homo sapien
32	150.5	22.6	292	4 Q00448	Q00448 homo sapien
33	149.5	22.4	214	11 Q9RQ08	Q9RQ08 mus musculu
34	148	22.2	178	11 Q91ZW9	Q91ZW9 mus musculu
35	148	22.2	208	11 Q91ZW7	Q91ZW7 mus musculu
36	148	22.2	1479	4 Q9UBG0	Q9UBG0 homo sapien
37	148	22.2	1479	4 Q9Y5P9	Q9Y5P9 homo sapien
38	148	22.2	1479	11 Q64449	Q64449 mus musculu
39	146.5	22.0	263	4 Q96Q03	Q96Q03 homo sapien
40	146.5	22.0	332	4 Q96Q05	Q96Q05 homo sapien
41	146.5	22.0	376	4 Q9BX53	Q9BX53 homo sapien
42	146.5	22.0	398	4 Q96Q08	Q96Q08 homo sapien
43	146.5	22.0	399	4 Q9H2X3	Q9H2X3 homo sapien
44	146	21.9	295	11 Q91ZW4	Q91ZW4 mus musculu
45	146	21.9	311	11 Q9DBV4	Q9DBV4 mus musculu

ALIGNMENTS

RESULT	ID	Q9UJ43	PRELIMINARY:	PRT:	385 AA.
AC	Q9UJ43:	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	L-SELECTIN PRECURSOR.				
GN	L-SELECTIN.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-HEMATOPOIETIC (B LYMPHOCYTE);				
RA	Flieger C.B.;				
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-HEMATOPOIETIC (B LYMPHOCYTE);				
RA	Flieger C.B.;				
RL	Thesis (1998), Freie Universitaet Berlin, Fachbereich Chemie.				
DR	EMBL; AJ246000; CAB55488.1; -.				
DR	HSSP; P14151; 1KJB.				
DR	InterPro; IPR000561; EGF-like.				
DR	InterPro; IPR001304; lectin_c.				
DR	InterPro; IPR002396; selectin.				
DR	InterPro; IPR000436; Sush1_SCR_CCP.				
DR	Pfam; PF00008; EGF; 1.				
DR	Pfam; PF00059; lectin_c; 1.				
DR	Pfam; PF00084; sush1; 2.				
DR	PRINTS; PR00343; SELECTIN.				
DR	SMART; SM00032; CCP; 2.				
DR	SMART; SM00034; CLECT; 1.				
DR	SMART; SM00181; EGF; 1.				
DR	PROSITE; PS00615; C-type-LECTIN_1; 1.				
DR	PROSITE; PS00641; C-type-LECTIN_2; 1.				
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.				

DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein; lectin; selectin; signal.
FT SIGNAL 1 51 POTENTIAL.
FT CHAIN 52 385 L-SELECTIN.
SQ SEQUENCE 385 AA; 43617 MW; 1205F691BA638EF1 CRC64;

Query Match 100.0%; Score 666; DB 4; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.8e-59;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNWMORARFCNDNTDVAIONKAEIYELEKTLPPSRSYWIGIRKIGIW 60
DB 52 WTYHYSEKPMNWMORARFCNDNTDVAIONKAEIYELEKTLPPSRSYWIGIRKIGIW 111

OY 61 TWGTTNKSILTEEAENMGDGEPPNKKNEKDCVEIYIKRNKDAGKWNDDACHKKLAALC 117
DB 112 TWGTTNKSILTEEAENMGDGEPPNKKNEKDCVEIYIKRNKDAGKWNDDACHKKLAALC 168

RESULT 2

ID 028629 PRELIMINARY; PRT; 376 AA.

AC 028629;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE L-SELECTIN PRECURSOR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY CORTEX;
RA Qian J., Marks R.M.;
RT "cDNA for rabbit L-selectin."
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U6535; AA67896.1; -
DR HSSP; P14151; 1KJB.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001304; lectin_c.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 2.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00343; CCP; 2.
DR SMART; SM00344; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW EGF-like domain; glycoprotein; lectin; selectin; signal.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 376 L-SELECTIN.
SQ SEQUENCE 376 AA; 42346 MW; 59F6AD530F490947 CRC64;

Query Match 89.8%; Score 598; DB 6; Length 376;
Best Local Similarity 88.0%; Pred. No. 1.3e-52;
Matches 103; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNWMORARFCNDNTDVAIONKAEIYELEKTLPPSRSYWIGIRKIGIW 60
DB 39 WTYHYSEKPMNWMORARFCNDNTDVAIONKAEIYELEKTLPPSRSYWIGIRKIGIW 98

OY 61 TWGTTNKSILTEEAENMGDGEPPNKKNEKDCVEIYIKRNKDAGKWNDDACHKKLAALC 117
DB 99 TWGTTNKSILTEEAENMGDGEPPNKKNEKDCVEIYIKRNKDAGKWNDDACHKKLAALC 155

RESULT 3
ID 063762 PRELIMINARY; PRT; 372 AA.

AC 063762;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE LYMPHOCYTE MEMBRANE PROTEIN A.11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=95369821; PubMed=7543874;
RA Sackstein R., Meng L., Xu X.M., Chin Y.H.;
RT "Evidence of post-transcriptional regulation of L-selectin gene
expression in rat lymphoid cells."
RL Immunology 85:198-204(1995).
DR EMBL; S79523; AAC60710.2; -
DR HSSP; P14151; 1KJB.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001304; lectin_c.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 2.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00343; CCP; 2.
DR SMART; SM00344; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein.
SQ SEQUENCE 372 AA; 42471 MW; 7DFD125610DD664A CRC64;

Query Match 83.6%; Score 557; DB 11; Length 372;
Best Local Similarity 81.2%; Pred. No. 1.9e-48;
Matches 95; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNWMORARFCNDNTDVAIONKAEIYELEKTLPPSRSYWIGIRKIGIW 60
DB 39 WTYHYSEKPMNWMORARFCNDNTDVAIONKAEIYELEKTLPPSRSYWIGIRKIGIW 98

OY 61 TWGTTNKSILTEEAENMGDGEPPNKKNEKDCVEIYIKRNKDAGKWNDDACHKKLAALC 117
DB 99 TWGTTNKSILTEEAENMGDGEPPNKKNEKDCVEIYIKRNKDAGKWNDDACHKKLAALC 155

RESULT 4
ID 095509 PRELIMINARY; PRT; 616 AA.

AC 095509;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE D780M13.1.3 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN CD62, GMP140)) (ISOFORM 3) (FRAGMENT).
GN SELP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.


```
DR EMBL: AL022146; CAA18144.1; -.
DR HSSP: P16109; 1KJD.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sush1_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sush1_6.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 6.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
DR EGF-like domain; glycoprotein.
KW EGF-like domain; glycoprotein.
FT NON_TER 616
SQ SEQUENCE 616 AA; 67736 MW; 35CD4BFADE61D724 CRC64;

Query Match
Best Local Similarity 66.8%; Score 445; DB 4; Length 616;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMWQRRARRCRDNYTDVAIQNKAEELEYEKLTPFSRSYYWIGIRKIGSIW 60
DB 42 WTYHSTAYKSWNISRKQCNRYTDVAIQNKNEDIDYLNKVLPIYSSTYYWIGIRKNNKTW 101
OY 61 TWGTFNKSUTEAEENWGDEPNKKKEDCEYIYIKRNKDAGKWNDDACHRLKALC 117
DB 102 TWGTFKALTNDEAEWADNPNKRNEDCEYIYIKSPSAPGKNWDECHLKKHALC 158

RESULT 5
OY 095507 PRELIMINARY; PRT; 740 AA.
AC 095507;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE D1780M13.1.1 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN
DE CD62, GMP140)) (ISOFORM 1) (FRAGMENT).
GN SELP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022146; CAA18142.1; -.
DR HSSP: P16109; 1KJD.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sush1_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sush1_8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein.
FT NON_TER 740
SQ SEQUENCE 740 AA; 81095 MW; 3B5F70A45B1A3CD4 CRC64;
```

```
Query Match
Best Local Similarity 66.8%; Score 445; DB 4; Length 740;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMWQRRARRCRDNYTDVAIQNKAEELEYEKLTPFSRSYYWIGIRKIGSIW 60
DB 42 WTYHSTAYKSWNISRKQCNRYTDVAIQNKNEDIDYLNKVLPIYSSTYYWIGIRKNNKTW 101
OY 61 TWGTFNKSUTEAEENWGDEPNKKKEDCEYIYIKRNKDAGKWNDDACHRLKALC 117
DB 102 TWGTFKALTNDEAEWADNPNKRNEDCEYIYIKSPSAPGKNWDECHLKKHALC 158

RESULT 6
OY 095508 PRELIMINARY; PRT; 740 AA.
AC 095508;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE D1780M13.1.2 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN
DE CD62, GMP140)) (ISOFORM 2) (FRAGMENT).
GN SELP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022146; CAA18143.1; -.
DR HSSP: P16109; 1KJD.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sush1_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sush1_8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein.
FT NON_TER 740
SQ SEQUENCE 740 AA; 81390 MW; ID2E35E6D93745CE CRC64;

Query Match
Best Local Similarity 66.8%; Score 445; DB 4; Length 740;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMWQRRARRCRDNYTDVAIQNKAEELEYEKLTPFSRSYYWIGIRKIGSIW 60
DB 42 WTYHSTAYKSWNISRKQCNRYTDVAIQNKNEDIDYLNKVLPIYSSTYYWIGIRKNNKTW 101
OY 61 TWGTFNKSUTEAEENWGDEPNKKKEDCEYIYIKRNKDAGKWNDDACHRLKALC 117
DB 102 TWGTFKALTNDEAEWADNPNKRNEDCEYIYIKSPSAPGKNWDECHLKKHALC 158

RESULT 7
OY 0951G2 PRELIMINARY; PRT; 484 AA.
AC 0951G2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
```


RT "The intron-exon structure of the porcine E-selectin-encoding gene.
 RL Gene 176:67-72(1996)
 DR EMBL: U37521; AAC46880.1; -.
 DR HSSP: P16581; IESL.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001304; lectin_C.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushl_SCR_CCP.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_C; 1.
 DR Pfam: PF00084; sushl; 4.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP; 4.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF; 1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01166; EGF_2; 1.
 DR EGF-like domain; Glycoprotein; lectin; Selectin.
 CO SEQUENCE 482 AA; 52341 MW; 97DC5D70BF115944 CRC64;

	Query Match	63.8%	Score 425;	DB 6;	Length 482;	
	Best Local Similarity	63.2%;	Pred. No. 6;le-35;			
	Matches	74;	Conservative	13;	Mismatches	35;
					Indels	0;
					Gaps	0;
OY	1 WTYHSEKPMNQRRRCRDNYITDVLVAIQNAELIEYLEKTLPFSSTYYIGIRKIGTW	60				
	: : : : : : : : : : : : : : : : : : :					
Db	23 WSYSTETMTFFDDASACCOORYTLHLVAIQNAEILEYELNSTFNYSASYWIGIRKINGTW	82				
	: : : : : : : : : : : : : : : : : : :					
OY	61 TWVGNKSLTEAEAWMGDEPNNKKNEECVEIYIKRNKDACKMWDNDCAHTKALAC	117				
	: : : : : : : : : : : : : : : : : : :					
Db	83 TWIGKKALTPEATYWADGEPNNKSNNECCVEIYIKRDDSCKMWDERSCSKKLALC	139				
	: : : : : : : : : : : : : : : : : : :					

RESULT	11			
028657				
ID	Q28657	PRELIMINARY;	PRT:	649 AA.
AC	Q28657;			
DT	01-NOV-1996	(TEMBLrel. 01, Created)		
DT	01-NOV-1996	(TEMBLrel. 01, last sequence update)		
DT	01-DEC-2001	(TEMBLrel. 19, last annotation update)		
DE	P-SELECTIN.			
OC	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
NC	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Vora D.K., Fang Z., Liva S.M., Parhaml F., Watson A.D., Drake T.A.,			
RA	Terillo M.C., Berliner J.A.;			
RT	"Induction of p-selectin by MM-LDL and its role in human			
RT	atherosclerosis".			
RL	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Warden C.H.;			
RL	Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; U39446; AAAB1385.1; ..			
DR	HSSP; P16109; IFSB.			
DR	InterPro; IPR000561; EGF_1like.			
DR	InterPro; IPR001304; Lectin_C.			
DR	InterPro; IPR002396; Selectin.			
DR	InterPro; IPR000436; Sushi_SCR_CCP.			
DR	Pfam; PF00008; EGF_1.			
DR	Pfam; PF00059; Lectin_C; 1.			
DR	Pfam; PF00084; sushi; 6.			
DR	PRINTS; PR00343; SELECTIN.			
DR	SMART; SM00032; CCP; 6.			
DR	SMART; SM00034; CLECT; 1.			
DR	SMART; SM00181; EGF; 1.			
DR	PROSITE; PS00615; C_Type_Lectin_1; 1.			

DR PROSITE, PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE, PS00022; EGF_1; UNKNOWN_1.
DR PROSITE, PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein; Lectin; Selectin.
SQ SEQUENCE 649 AA; 71755 MW; ECCDC8647B84DC31 CRR64

Query Match	63.1%	Score 420;	DB 6;	Length 649;
Best Local Similarity	65.0%;	Pred. No. 2.8e-34;		
Matches 78;	Conservative 11;	Mismatches 25;	Indels 6;	Gaps 3
Oy	1	WYHYHSEKPMNQRRARPCFRDNYTDLVAQNKAETLEYLEKTLPEFSRYTWIGIRK---	IG 57	
		: : : : : : :		
Db	42	WYHYHSNKTYSNVNSAFCQKYTTDLVAQNNEIDYLIETIPYYSYWIGIRKQDQIN	101	
		: : : : : : :		
Oy	58	GWTWVAGTNGSLTTEAEENNGDCEPNNKKKKKKEPCVEIYIKRNNDACGWNDDACGKTLKAALC	117	
		: : : : : : :		
Db	102	G---TWVG--KNTLTEEEENADNEPNNKRNNDQCVETIYKLSLAPGCKWMDPECKRRALC	158	
		: : : : : : :		

RESULT	12		
09GLF0		PRELIMINARY:	PRT: 609 AA.
ID	09GLF0		
AC	09GLF0;		
DT	01-MAR-2001 (Tremblrel, 16, Created)		
DT	01-MAR-2001 (Tremblrel, 16, Last sequence update)		
DT	01-DEC-2001 (Tremblrel, 19, Last annotation update)		
DE	E-SELECTIN.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_TaxID=9615;		
NP	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-ENDOTHELIAL CELLS.		
RA	Zhang L., Shi Y., Wu H., Zhang G.;		
RT	"Cloning and sequencing of beagle E-selectin genomic DNA and		
RT	comparison with other species."		
RL	Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.		
DR	HML: AF87257; ANG10039.1; -		
DR	HSP: p16581; 1ESL.		
DR	InterPro: IPR000561; EGF-like.		
DR	InterPro: IPR001304; lectin_c.		
DR	InterPro: IPR002396; Selectin.		
DR	InterPro: IPR000436; Sushi_SCR_CCP.		
DR	Pfam: PF00059; lectin_c; 1.		
DR	Pfam: PF00084; sushi_6.		
DR	PRINTS; PRO0343; SELECTIN.		
DR	SMART; SM00032; CCP; 6.		
DR	SMART; SM00034; CLECT; 1.		
DR	SMART; SM00181; EGF; 4.		
DR	SMART; SM00001; EGF-like; 1.		
DR	PROSITE; PS50041; C_type_lectin_2; 1.		
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.		
DR	lectin; Selectin.		
QW	SEQUENCE 609 AA: 66073 MW; 41B62D1F4D23881F CRC64;		

	Query Match	61.3%	Score 408;	DB 6;	Length 609;	
	Best Local Similarity	59.0%;	Pred. No. 4.3e-33;			
	Matches	69;	Conservative	18;	Mismatches	30; Indels 0; Gaps 0.
Qy	1	WYHYSEKPMNQARARECDNYDVLVAIQNAEIEYLEKTLPFSRXYWIGIRKGIM	60			
Dd	23	MSYNSTEMTMTDEASTYTCQQRTHLVALIQNEELIKLYNSMSTPTTYTWIGIRKNKM	82			
Qy	61	TWVGINKSLTEAEAWMGDGEPNNKKNKEDCBEIYTRNNDACKWMNDACHKLKAALC	117			
Dd	83	TWIGTKLLTEAKMWAPGEPNNKQNDEDCBEIYTRPDSDSKWMNDERCKKLKALC	139			
RESULT	13					
	229097					

QY	DB	SEQUENCE	Score	DB 6;	Length	610;
QY	1	WYHYSEKPMNQARFPCRDNYTDVAIONKAEIYELEKTPFSSRYWIGIRKIGIW 60	58.9%	392.5;	DB 6;	Length 610;
DB	22	WYSASTTWMFPDEASAYCQQRTHVLAIONOEIKYLSTNHSPPYWIGIRKYNKW 81	Best Local Similarity 60.7%; Pred. No. 1.6e-31;			
		Matches 71; Conservative 12; Mismatches 33; Indels 1; Gaps 1				
QY	61	TWVGTKSLTEAEWKGDESPNNKKKECCVEIYIKRNKADGWNDACHIKALC 117				
DB	82	WVGTKPLTEAKWAPDEPNKKON-EDCVEIYIKRYKADGWNDENCKRRKALC 137				
RESULT	15					
ID	Q9UJ71	PRELIMINARY; PRT; 328 AA.				
AC	Q9UJ71.					
DT	01-MAY-2000 (TREMBLrel. 13, Created)					
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)					
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)					
DE	LANGERIN.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.					
OX	NCBI_TaxID=9606;					
RA	SEQUENCE FROM N.A.					
RA	Valladeau J., Revel O., Deutter-Dambuyant C., Moore K., Kleijmeer M.,					
RA	Duvert-Francis V., Vincent C., Schmitt D., Davoust J., Caux C.,					
RA	Lebecque S., Saeland S.;					
RT	"Langerin, a new transmembrane C-type lectin specific to Langerhans					
RT	cells, induces the formation of Birbeck granules.";					
RL	Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.					
DR	EMBL; AJ242859; CAB62403.1; -					
DR	HSSP; P20693; IHLJ					
DR	InterPro; IPR001304; lectin_c.					
DR	Pfam; PF00059; lectin_c; 1.					
DR	SMART; SM00034; CLECT; 1.					
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.					
DR	PROSITE; PS00041; C_TYPE_LECTIN_2; 1.					
SO	SEQUENCE 328 AA; 36697 MW; 83DF5A1B347D1B62 CRC64;					
Query Match	26.6%; Score 177; DB 4; Length 328;					
Best Local Similarity	34.7%; Pred. No. 5.3e-10;					
Matches 41; Conservative 16; Mismatches 53; Indels 8; Gaps 4;						
QY	3	YHYSEKPMNQARFPCRDNYTDVAIONKAEIYELEKTPFSSRYWIGIRKIGIW 60				
DB	207	YHYSLIPTWISAQFCRSNSHLTSVISEQSEFLVTA--GGILYIIGLTGKMGMDW 264				
QY	61	TWVGTKSLTEAEWKGDESPNNKKKECCVEIYIKRNKADGWNDACHIKALC 117				
DB	265	SWVDTFPNKQYASRFWIPGEPNNAAGNNEHGN1---KAPSLQAMNADPCDKTFLIC 319				
RESULT	16					
ID	Q9VOX3	PRELIMINARY; PRT; 359 AA.				
ID	Q9VOX3					
DT	01-MAY-2000 (TREMBLrel. 13, Created)					

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE CG3958 PROTEIN (H105328P).
 CN LECTIN-24DB OR CG2958.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandal D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadien E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Fjosek C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Foster J.M., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jatala M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 RA Klambo B.E., Kodra C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Stryckas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y, CN BW SP;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Fartan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO03577; AFS1038.1; -;
 DR EMBL: AY061025; AAL28573.1; -;
 DR HSSP: P14151; 1KJB
 DR P1Ybase: F8gn0040102; lectin-24db.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PSS00041; C-TYPE LECTIN.2; 1.
 SO SEQUENCE 359 AA; 41546 MW; 1895E525594584C4 CRC64;

Query Match 24.6%; Score 166; DB 5; Length 359;
 Best Local Similarity 33.6%; Pred. No. 7.6e-09;

Matches 36; Conservative 16; Mismatches 47; Indels 8; Gaps 2;
 QY 11 NMQRARRCRDNYTDVLAIONKAELEYEKLTPFSRSYWGIRKIGGWTWVGKSLT 70
 DB 255 DWQSAVDPCRWGXYIAIKDDELDIAISAR--DDKSYWGINDLQSSNNYVSASGRE 312
 QY 71 EAEWNGGDEPPNNKKKEDCEIYIKRKKDKAGWDDACHKAKALC 117
 DB 313 VEFLLMNMNGEPHNGEDNCVELT-----RSKMNDDPCRRKKHYIC 353
 RESULT 17
 0951AB PRELIMINARY; PRT; 381 AA.
 AC 0951AB;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE DENDRITIC CELL-SPECIFIC ICAM-3 GRABING NONINTEGRIN.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21465051; PubMed=11581396;
 RA Baribud F., Pohlmann S., Sparwasser T., Kimata M.T., Choi Y.K.,
 RA Haggarty B.S., Ahmed N., Macfarlan T., Edwards T.G., Leslie G.J.,
 RA Arason J., Reinhardt T.A., Kimata J.T., Littman D.R., Hoxie J.A.,
 RA Doms R.W.;
 RT "Functional and antigenic characterization of human, rhesus macaque,
 RT pigtailed macaque, and murine dc-sign.";
 RL J. Virol. 75:10281-10289(2001).
 DR EMBL: AF369755; AAL14438.1; -;
 SO SEQUENCE 381 AA; 42897 MW; 01E7B0B42C91D49 CRC64;
 Query Match 24.4%; Score 162.5; DB 6; Length 381;
 Best Local Similarity 29.6%; Pred. No. 1.9e-08;
 Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;
 QY 1 WT-----YHSEKPMNQRRARRCRDNYTDVLAIONKAELEYEKLTPFSRSYWGIR 54
 DB 237 WFFQGNCTFWSNSORNNHNSITACQEGVQADLVYKSAEONFLOLQSSRNRTWGLS 296
 QY 55 KIG--GIWTWGTNKSLEAEENMGDEPPNNKKKEDCEIYIKRKKDKAGWDDACHKL 112
 DB 297 DLNHEGTWQWDGSLPLPFRQYNNKGEPPN-IGEDCAEP-----SGNGMNDKCNLA 349
 QY 113 KALC 117
 DB 350 KFWIC 354
 RESULT 18
 0960Q3 PRELIMINARY; PRT; 268 AA.
 AC 0960Q3;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE SDC-STGNIA TYPE III ISOFORM.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=11337487;
 RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,

RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
 RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts.";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042227; AAK91852.1;
 SO SEQUENCE 268 AA; 30427 MW; 113E1ED1B5748693 CRC64;

Query Match 24.2%; Score 161.5; DB 4; Length 268;
 Best Local Similarity 29.6%; Pred. No. 1.5e-08;
 Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;

OY 1 WT-----YHYSEKPMNMQARARFCRDNYTDLVAIQNKAIELEYLKTLPFSRSYYWIGIR 54
 DB 124 WTEFGNCYFMSNSORNMWDSITACKEVGAQLVYIKSAEQNFLQLOSSRSNRFTWIGLS 183
 OY 55 KIG--GIWTVGNTKSLTEAEENMGDEPNKKKNECDVEIYIKRNKDGAKNDACAKL 112
 DB 184 DLNQGWTQWVDGSPLLPSFKQYWNKGPNN-VGEEDCAEF-----SGNGWMDCKNLA 236
 OY 113 KAALC 117
 DB 237 KFWIC 241

RESULT 19
 ID 096007 PRELIMINARY; PRT; 312 AA.
 AC 096007;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE MDC-SIGNIA TYPE III ISOFORM.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed-11337487;
 RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
 RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts.";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042223; AAK91848.1;
 SO SEQUENCE 312 AA; 35178 MW; 95761618EF36AFBE CRC64;

Query Match 24.2%; Score 161.5; DB 4; Length 312;
 Best Local Similarity 29.6%; Pred. No. 1.8e-08;
 Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;

OY 1 WT-----YHYSEKPMNMQARARFCRDNYTDLVAIQNKAIELEYLKTLPFSRSYYWIGIR 54
 DB 168 WTEFGNCYFMSNSORNMWDSITACKEVGAQLVYIKSAEQNFLQLOSSRSNRFTWIGLS 227
 OY 55 KIG--GIWTVGNTKSLTEAEENMGDEPNKKKNECDVEIYIKRNKDGAKNDACAKL 112
 DB 228 DLNQGWTQWVDGSPLLPSFKQYWNKGPNN-VGEEDCAEF-----SGNGWMDCKNLA 280
 OY 113 KAALC 117
 DB 281 KFWIC 285

RESULT 20
 0960P9

ID 0960P9 PRELIMINARY; PRT; 334 AA.
 AC 0960P9;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE SDC-SIGNIB TYPE II ISOFORM.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed-11337487;
 RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
 RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts.";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042231; AAK91856.1;
 SO SEQUENCE 334 AA; 37843 MW; 8E796FC4111C86BF CRC64;

Query Match 24.2%; Score 161.5; DB 4; Length 334;
 Best Local Similarity 29.6%; Pred. No. 2e-08;
 Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;

OY 1 WT-----YHYSEKPMNMQARARFCRDNYTDLVAIQNKAIELEYLKTLPFSRSYYWIGIR 54
 DB 190 WTEFGNCYFMSNSORNMWDSITACKEVGAQLVYIKSAEQNFLQLOSSRSNRFTWIGLS 249
 OY 55 KIG--GIWTVGNTKSLTEAEENMGDEPNKKKNECDVEIYIKRNKDGAKNDACAKL 112
 DB 250 DLNQGWTQWVDGSPLLPSFKQYWNKGPNN-VGEEDCAEF-----SGNGWMDCKNLA 302
 OY 113 KAALC 117
 DB 303 KFWIC 307

RESULT 21
 ID 0960Q4 PRELIMINARY; PRT; 360 AA.
 AC 0960Q4;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE SDC-SIGNIA TYPE II ISOFORM.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed-11337487;
 RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
 RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts.";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042226; AAK91851.1;
 SO SEQUENCE 360 AA; 41009 MW; 6ABE2B9AEDFPAAB8 CRC64;

Query Match 24.2%; Score 161.5; DB 4; Length 360;
 Best Local Similarity 29.6%; Pred. No. 2.2e-08;
 Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;

```

OY 1 WT-----YHYSEKPMNQARARPCRDNYTDLVAIONKAIEIYLEKTLPEFSRSYWIIGIR 54
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 WTFPGNGCYFMSNSQRNMHDSITACKKEYGAOLVWIKSAEONFLOLQSSRSNRTFMGILS 275
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 55 KIG--GIWTWGTGNTSLTEEAENMGDGEFENPNKKNEDEVEIYIKRNKDGAKYNDACHTL 112
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 276 DLNQGTCQWMDGSPFLPSFKQYWNRGEPNN--VGEEDCAEF-----SGNGNNDKCNLA 328
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 113 KAALC 117
      | | |
DB 329 KFWIC 333

RESULT 22
O96005 PRELIMINARY; PRT; 380 AA.
O96005
AC O96005;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SDC-SIGN1A TYPE I ISOFORM.
GN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;

Query Match 24.2%; Score 161.5; DB 4; Length 380;
Best Local Similarity 29.6%; Pred. No. 2,3e-08;
Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4.

OY 1 WT-----YHYSEKPMNQARARPCRDNYTDLVAIONKAIEIYLEKTLPEFSRSYWIIGIR 54
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 236 WTFPGNGCYFMSNSQRNMHDSITACKKEYGAOLVWIKSAEONFLOLQSSRSNRTFMGILS 295
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 55 KIG--GIWTWGTGNTSLTEEAENMGDGEFENPNKKNEDEVEIYIKRNKDGAKYNDACHTL 112
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 296 DLNQGTCQWMDGSPFLPSFKQYWNRGEPNN--VGEEDCAEF-----SGNGNNDKCNLA 348
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 113 KAALC 117
      | | |
DB 349 KFWIC 353

RESULT 23
O96000 PRELIMINARY; PRT; 380 AA.
O96000;
AC O96000;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SDC-SIGN1B TYPE I ISOFORM.
GN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;

```

```

RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.:
RA "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT specific ICAM-3-grabbing NonIntegrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL; AF042230; AAK91855.1; -.
SQ SEQUENCE 380 AA: 43125 MW; F3D098F9FB7D044B CRC64;

Query Match 24.2%; Score 161.5; DB 4; Length 380;
Best Local Similarity 29.6%; Pred. No. 2.3e-08;
Matches 37; Conservative 16; Mismatches 37; Indels 15; Gaps

OY 1 WT-----YHYSEKPMNWQRRARFCRDNTDVAIONKAEIYLEKTLTFFSRSYWIGIR 54
DB 236 WFFPGQNCYFMSNSQNMWHDSTACQEVGAQLVYIKSAEQNFLOQSSRSNRFWMGIS 295
OY 55 KIG--GIWTWVGTSNLSLTEAEWNGDEFPNNKKNEDEVEIYIKRNKDGAKNDDACHL 112
DB 296 DINEGTQWQVDSPLLPFKQYMNKGEPNN--VGEDCAEF-----SGNGWMDKC NLA 348
OY 113 KAALC 117
DB 349 KFWIC 353

RESULT 24
OY 095LC6 PRELIMINARY; PRT; 381 AA.
AC 095LC6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DI 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DT DENDRITIC CELL-SPECIFIC ICAM-3 GRABBIN NONINTEGRIN.
OS Macaca nemestrina (Pig-tailed macaque).
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC NCBI_TaxId=9545; Macaca.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21465051; PubMed=11581396;
RX Barbaud F., Pohlmann S., Sparwasser T., Kimata M.T., Choi Y.K.,
RX Haggarty B.S., Ahmad N., Macfarlan T., Edwards T.G., Leslie G.J.,
RX Arnson J., Reinhart T.A., Kimata J.T., Littman D.R., Hoxie J.A.,
RA Doms R.W.:
RA "Functional and antigenic characterization of human, rhesus macaque,
RT pigtailed macaque, and murine dc-sgn.";
RL J. Virol. 75:10281-10289(2001).
DR EMBL; AF343727; AALL4428.1; -.
KW Integrin.
SQ SEQUENCE 381 AA: 42951 MW; CAF6E2E3D454B74A CRC64;

Query Match 24.2%; Score 161.5; DB 6; Length 381;
Best Local Similarity 29.6%; Pred. No. 2.3e-08;
Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps

OY 1 WT-----YHYSEKPMNWQRRARFCRDNTDVAIONKAEIYLEKTLTFFSRSYWIGIR 54
DB 237 WFFPGQNCYFMSNSQNMWHDSTACQEVGAQLVYIKSAEQNFLOQSSRSNRFWMGIS 296
OY 55 KIG--GIWTWVGTSNLSLTEAEWNGDEFPNNKKNEDEVEIYIKRNKDGAKNDDACHL 112
DB 297 DINEGTQWQVDSPLLPFKQYMNKGEPNN--VGEDCAEF-----SGNGWMDKC NLA 349
OY 113 KAALC 117
DB 350 KFWIC 354

```

```
RESULT 25
ID Q9NNX6 PRELIMINARY; PRT; 404 AA.
AC Q9NNX6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MEMBRANE-ASSOCIATED LECTIN TYPE-C (PROBABLE MANNOSYL-BINDING C-TYPE
DE LECTIN DC-SIGN) (MDC-SIGNIA TYPE I ISOFORM).
GN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=92390446; PubMed=1518669;
RA Curtis B.M., Scharnowske S., Watson A.J.;
RT "Sequence and expression of a membrane-associated C-type lectin that
RT exhibits CD4-independent binding of human immunodeficiency virus
RT envelope glycoprotein gp 120."
RL Proc. Natl. Acad. Sci. U.S.A. 89:8356-8360(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20432267; PubMed=10975799;
RA Solleux E.J., Barten R., Trowsdale J.;
RT "DC-SIGN, a related gene, DC-SIGNR, and CD23 form a cluster on
RT 19p13."
RL J. Immunol. 165:2937-2942(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21157496; PubMed=11257134;
RA Bashirova A.A., Geijtenbeek T.B.H., van Duijnhoven G.C.F.,
RA van Vliet S.J., Eilander J.B.G., Martin M.P., Wu L., Martin T.D.,
RA Viebig N., Knolle P.A., KewalRamani V.N., van Kooyk Y., Carrington M.;
RT "A dendritic cell-specific intercellular adhesion molecule 3-grabbing
RT nonintegrin (DC-SIGN)-related protein is highly expressed on human
RT liver sinusoidal endothelial cells and promotes HIV-1 infection."
RL J. Exp. Med. 193:671-678(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: M98457; AAF7072.1; -
DR EMBL: AF209479; AAG13814.1; -
DR EMBL: AF290886; AAK20997.1; -
DR EMBL: AY042221; AAK91846.1; -
DR HSSP: P20693; IHLI.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; UNKNOMN_1.
DR PROSITE: PS00641; C_TYPE_LLECTIN_2; 1.
KM lectin.
SQ SEQUENCE 404 AA; 45775 MW; A23FA246014533C0 CRC64;
```

Query Match 24.2%; Score 161.5; DB 4; Length 404;
Best Local Similarity 29.6%; Pred. No. 2.5e-08;

Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;

```
QY 1 WT-----YHSEKPMNMOARRFCDNYTDVAIQNKAIEYLEKTLPFSSRSYWGIR 54
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 260 WTFPGNCYFMSNSQRNWDSTACKEVGAQLVYIKSAEONFLQSSRSNFTWMGIS 319
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 55 KIG--GIWTWGTNKSLTAEENWGDGEPNNKKNKEDCEYIYIKRNKDGKWNDDACHKL 112
```

```
DB 320 DLNQEGTWMVDGSPLLPSEFKQYWNRGEPNN-VGEEDCAEF-----SGNGWDDCKNLA 372
QY 113 KALC 117
DB 373 KEWIC 377
```

```
RESULT 26
ID Q96001 PRELIMINARY; PRT; 404 AA.
AC Q96001;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MDC-SIGNIB TYPE I ISOFORM.
GN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042229; AAK91854.1; -
SQ SEQUENCE 404 AA; 45570 MW; D4D570727D9DCC30 CRC64;
```

Query Match 24.2%; Score 161.5; DB 4; Length 404;
Best Local Similarity 29.6%; Pred. No. 2.5e-08;

Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;

```
QY 1 WT-----YHSEKPMNMOARRFCDNYTDVAIQNKAIEYLEKTLPFSSRSYWGIR 54
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 260 WTFPGNCYFMSNSQRNWDSTACKEVGAQLVYIKSAEONFLQSSRSNFTWMGIS 319
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 55 KIG--GIWTWGTNKSLTAEENWGDGEPNNKKNKEDCEYIYIKRNKDGKWNDDACHKL 112
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 320 DLNQEGTWMVDGSPLLPSEFKQYWNRGEPNN-VGEEDCAEF-----SGNGWDDCKNLA 372
QY 113 KALC 117
DB 373 KEWIC 377
```

```
RESULT 27
ID Q14538 PRELIMINARY; PRT; 292 AA.
AC Q14538;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MACROPHAGE LECTIN 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96108853; PubMed=8598452;
RA Suzuki N., Yamamoto K., Toyoshima S., Osawa T., Irimura T.;
RT "Molecular cloning and expression of cDNA encoding human macrophage c-
RT type lectin: its unique carbohydrate binding specificity for Tn
RT antigen."
RL J. Immunol. 156:128-135(1996).
```

```
DR EMBL: D50532; BAA09101.1; -.
```



```

RESULT 31
003969
ID 003969 PRELIMINARY; PRT; 287 AA.
AC 003969;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DR HEPATIC LECTIN H2, VARIANT (ASIALOGLYCOPROTEIN RECEPTOR H2).
GN L-H2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92184202; PubMed=1371982;
RA Paletta E., Racevskis J., Stockert R.J.;
RT "Differences in the abundance of variably spliced transcripts for the
RT second asialoglycoprotein receptor polypeptide, H2, in normal and
RT transformed human liver.";
RL Hepatology 15:395-402(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=HEPATOMA;
RA Strausberg R.;
RT Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SUBUNIT: THE FUNCTIONING LIGAND-BINDING UNIT OF THIS RECEPTOR IS
CC THOUGHT TO BE AT LEAST A DIMER.
CC -1- MISCELLANEOUS: THIS PROTEIN IS A VARIANT OF HUMAN HEPATIC LECTIN
CC H2 (AC P07307).
CC -1- MISCELLANEOUS: THE ASIALOGLYCOPROTEIN RECEPTOR IS EXPRESSED
CC EXCLUSIVELY IN HEPATIC PARENCHYMAL CELLS IN MAMMALS. AFTER LIGAND
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC DIASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC SURFACE.
CC -1- MISCELLANEOUS: AFTER REMOVAL OF STATIC ACID MONOMERS FROM THE
CC COMPLEX CARBOHYDRATE MOETIES OF PLASMA GLYCOPROTEINS, TERMINAL
CC GALACTOSE UNITS ARE RECOGNIZED BY THIS HEPATIC RECEPTOR.
CC -1- MISCELLANEOUS: THE USUAL ORIENTATION OF THIS PROTEIN ACROSS THE
CC MEMBRANE IS POSTULATED TO OCCUR BY RECOGNITION OF AN INTERNAL
CC TRANSMEMBRANE SEGMENT AS A SIGNAL SEQUENCE.
DR EMBL: X55283; CAA38997.1; -.
DR EMBL: BC017251; AAH17251.1; -.
DR HSSP; P06734; 1HTI.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF000059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS0041; C_TYPE_LLECTIN_2; 1.
KM Lectin, glycoprotein, Receptor, Endocytosis, Transmembrane.
FT CAROHD 78 78 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CAROHD 146 146 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CAROHD 281 281 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ SEQUENCE 287 AA; 32550 MW; E9528ACFCB1435 CRC64;

Query Match 22.6%; Score 150.5; DB 4; Length 287;
Best Local Similarity 30.3%; Pred. No. 2.2e-07;
Matches 37; Conservative 21; Mismatches 47; Indels 17; Gaps 6;

```

```

Db 275 VC 276
!
RESULT 32
000448
ID 000448 PRELIMINARY; PRT; 292 AA.
AC 000448;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DR ASIALOGLYCOPROTEIN RECEPTOR.
GN ASGPR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86016723; PubMed=3863106;
RA Spiess M., Lodish H.F.;
RT "Sequence of a second human asialoglycoprotein receptor: conservation
RT of two receptor genes during evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6465-6469(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92184202; PubMed=1371982;
RA Paletta E., Stockert R.J., Racevskis J.;
RT "Differences in the abundance of variably spliced transcripts for the
RT second asialoglycoprotein receptor polypeptide, H2, in normal and
RT transformed human liver.";
RL Hepatology 15:395-402(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee D.G., Lee S.G., Kim K.L., Hahn K.-S.;
RT "Sequences of cDNAs encoding two subunits of asialoglycoprotein
RT receptor in human liver.";
RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: U97197; AAB58308.1; -.
DR HSSP; P06734; 1HTI.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF000059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS0041; C_TYPE_LLECTIN_2; 1.
KM Receptor.
SQ SEQUENCE 292 AA; 33087 MW; 8DC85A2B9757CA9A CRC64;

Query Match 22.6%; Score 150.5; DB 4; Length 292;
Best Local Similarity 30.3%; Pred. No. 2.2e-07;
Matches 37; Conservative 21; Mismatches 47; Indels 17; Gaps 6;

```



```

RA Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;
RT "A urokinase receptor-associated protein with specific collagen-
RT binding properties.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:169-176(1998).
DR EMBL: AF107292; AAF14192.1; -.
DR EMBL: AB014609; BAA31684.1; -.
DR HSSP: P02751; 2FN2.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR000566; Lipocln_cyFABP.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00059; Lectin_C; 8.
DR PRINTS: PR00013; FNTYPEII.
DR PRODOM: PD000995; FN_Type_II; 1.
DR SMART: SM00034; CLECT; 8.
DR SMART: SM00458; RICIN; 1.
DR PROSITE: PS00615; C_Type_Lectin_1; 3.
DR PROSITE: PS50041; C_Type_Lectin_2; 8.
DR PROSITE: PS00023; FIBRONECTIN_2; UNKNOWN_1.
DR PROSITE: PS50213; LIPOCALIN; UNKNOWN_1.
DR PROSITE: PS50231; RICIN_B_Lectin; 1.
KW Receptor; Kinase
SQ SEQUENCE 1479 AA; 166654 MW; C7583EA78E2792D1 CRC64;

Query Match 22.2%; Score 148; DB 4; Length 1479;
Best Local Similarity 30.3%; Pred. No. 2.8e-06;
Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

OY 3 YHSEKPMNQRRRCRDNYTDVAIONKAELEYLEKTLPSRSRYWIGIR--KIGGIW 60
DB 394 YRLQAKRRSQESKACLRGGDLVSIHSMALERTTKOIKOEVEELWGLDLKLQNF 453
OY 61 TWV-TGNKSLTEAEENWGDEPN--KKNKDCVEIYIKRNKDGAKNNDACHKLKALC 117
DB 454 EWSDGSLVFT----HWHPPEPNFRDLSLEDVTTW----GPEGRNDSFCNQLSPSIC 504

RESULT 37
OY5P9 PRELIMINARY; PRT; 1479 AA.
AC OY5P9:
DT 01-NOV-1998 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ENDOCYTIC RECEPTOR ENDO180.
GN ENDO180.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20148849; PubMed=10683150;
RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;
RT "Endo180, an endocytic recycling glycoprotein related to the
RT macrophage mannose receptor is expressed on fibroblasts, endothelial
RT cells and macrophages and functions as a lectin receptor.";
RL J. Cell Sci. 113:1021-1032(2000).
EMBL: AF134838; AAD30280.1; -.

```

```

DR HSSP: P02751; 2FN2.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR000566; Lipocln_cyFABP.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00059; Lectin_C; 8.
DR Pfam: PF00652; Ricin_B_lectin; 1.
DR PRINTS: PR00013; FNTYPEII.
DR PRODOM: PD000995; FN_Type_II; 1.
DR SMART: SM00034; CLECT; 8.
DR SMART: SM00458; RICIN; 1.
DR PROSITE: PS00615; C_Type_Lectin_1; 3.
DR PROSITE: PS50041; C_Type_Lectin_2; 8.
DR PROSITE: PS00023; FIBRONECTIN_2; UNKNOWN_1.
DR PROSITE: PS50213; LIPOCALIN; UNKNOWN_1.
DR PROSITE: PS50231; RICIN_B_Lectin; 1.
KW Receptor.
SQ SEQUENCE 1479 AA; 166669 MW; 9F4BAF355F036FCE CRC64;

Query Match 22.2%; Score 148; DB 4; Length 1479;
Best Local Similarity 30.3%; Pred. No. 2.8e-06;
Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

OY 3 YHSEKPMNQRRRCRDNYTDVAIONKAELEYLEKTLPSRSRYWIGIR--KIGGIW 60
DB 394 YRLQAKRRSQESKACLRGGDLVSIHSMALERTTKOIKOEVEELWGLDLKLQNF 453
OY 61 TWV-TGNKSLTEAEENWGDEPN--KKNKDCVEIYIKRNKDGAKNNDACHKLKALC 117
DB 454 EWSDGSLVFT----HWHPPEPNFRDLSLEDVTTW----GPEGRNDSFCNQLSPSIC 504

RESULT 38
OY5P9 PRELIMINARY; PRT; 1479 AA.
AC OY5P9:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE LECTIN LAMBDA.
GN MRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355501; PubMed=8702911;
RA Wu K., Yuan J., Lasky L.A.;
RT "Characterization of a novel member of the macrophage mannose receptor
RT type C lectin family.";
RL J. Biol. Chem. 271:21323-21330(1996).
DR EMBL: U56734; AAC52729.1; -.
DR HSSP: P02751; 2FN2.
DR MGD: MGI:107818; Mrc2.
DR InterPro: IPR001128; Cyt_P450.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR000566; Lipocln_cyFABP.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00059; Lectin_C; 8.
DR PRINTS: PR00013; FNTYPEII.
DR PRODOM: PD000995; FN_Type_II; 1.
DR SMART: SM00034; CLECT; 8.
DR SMART: SM00458; RICIN; 1.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
DR PROSITE: PS00615; C_Type_Lectin_1; 3.
DR PROSITE: PS50041; C_Type_Lectin_2; 8.

```

DR PROSITE: PS00023; FIBRONECTIN_2; UNKNOWN_1.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE: PS50231; RICIN B LECTIN; 1.
SO SEQUENCE 1479 AA; 167112 MW; 620456E109B48C1 CRC64;

Query Match 22.0%; Score 148; DB 11; Length 1479;
Best Local Similarity 30.3%; Pred. No. 2.8e-06;
Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

OY 3 YHYSEKPMNQRRARPCRDNTDLVAIONKAEIYELEKTLPPSRSYWIGIR--NIGSIW 60
DB 393 YRLQAEKRSWQESKRACLRGGDLISHSMALEFRTQIKQEVBELWIGLNDLKLQNMFF 452
OY 61 TWV-GTNKSLTEEAENMGDGEPPN-KKNEKDCVEIYIKRNDAGKWNDDACHKLAALC 117
DB 453 EMSDGSLSVFT---HMHPEEPNFRDLSLDCVITW---GPEGKWNDSPCNQLSPISIC 503

RESULT 39
O960P3 PRELIMINARY; PRT; 263 AA.
AC O960P3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SOC-SIGN2 TYPE III ISOFORM.
CN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed-1137487;
RA Mummidi S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
RT Transcripts";
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042240; AAK91865.1;--
SO SEQUENCE 263 AA; 30102 MW; C611C762C50424CF CRC64;

Query Match 22.0%; Score 146.5; DB 4; Length 263;
Best Local Similarity 29.1%; Pred. No. 5e-07;
Matches 34; Conservative 12; Mismatches 56; Indels 15; Gaps 4;

OY 1 WT-----YHYSEKPMNQRRARPCRDNTDLVAIONKAEIYELEKTLPPSRSYWIGIR 54
DB 136 WTFPGNCFPMNSQNRNHDVTACQEVRAQLVYIKTAEQNFLOLQTSRSNRFSGMGLS 195
OY 55 KIG--GIWTWGTNKSLEEAENMGDGEPPNKKNEKDCVEIYIKRNDAGKWNDDAC 109
DB 196 DLNDEGTQWWDGSPSPSFQRYWNSGEPNNSGN-EDCAEF-----SGSGWMDNRC 245

RESULT 40
O960P5 PRELIMINARY; PRT; 332 AA.
AC O960P5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SOC-SIGN2 TYPE I ISOFORM.
CN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]

RP SEQUENCE FROM N.A.
RX PubMed-1137487;
RA Mummidi S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN
RT Transcripts";
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042238; AAK91863.1;--
SO SEQUENCE 332 AA; 37953 MW; CGDEF92C1B073C6 CRC64;

Query Match 22.0%; Score 146.5; DB 4; Length 332;
Best Local Similarity 29.1%; Pred. No. 6.6e-07;
Matches 34; Conservative 12; Mismatches 56; Indels 15; Gaps 4;

OY 1 WT-----YHYSEKPMNQRRARPCRDNTDLVAIONKAEIYELEKTLPPSRSYWIGIR 54
DB 205 WTFPGNCFPMNSQNRNHDVTACQEVRAQLVYIKTAEQNFLOLQTSRSNRFSGMGLS 264
OY 55 KIG--GIWTWGTNKSLEEAENMGDGEPPNKKNEKDCVEIYIKRNDAGKWNDDAC 109
DB 265 DLNDEGTQWWDGSPSPSFQRYWNSGEPNNSGN-EDCAEF-----SGSGWMDNRC 314

Search completed: September 7, 2002, 10:23:17
Job time: 545 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:47 : Search time 18.65 seconds
(without alignments)
242.905 Million cell updates/sec

Title: US-09-119-209-2_COPY_39_155

Perfect score: 666
Sequence: 1 WTYHYSEKMNQRRARFCR.....NKDAGKMNDACHKLKALIC 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	666	100.0	372	1	LEM1_HUMAN
2	666	100.0	372	1	LEM1_PANTR
3	662	99.4	372	1	LEM1_PONPY
4	631	94.7	372	1	LEM1_MACMU
5	631	94.7	372	1	LEM1_PAPHA
6	581	87.2	370	1	LEM1_BOVIN
7	557	83.6	372	1	LEM1_MOUSE
8	557	83.6	372	1	LEM1_RAT
9	464	69.7	769	1	LEM3_SHEEP
10	455	68.3	768	1	LEM3_MOUSE
11	451	67.7	646	1	LEM3_BOVIN
12	445	66.8	485	1	LEM2_BOVIN
13	445	66.8	830	1	LEM3_HUMAN
14	442	66.4	768	1	LEM3_RAT
15	434	65.2	551	1	LEM2_RABIT
16	425	63.8	484	1	LEM2_PIG
17	415	62.3	610	1	LEM2_HUMAN
18	408	61.3	549	1	LEM2_RAT
19	408	61.3	611	1	LEM2_CANFA
20	405	60.8	612	1	LEM2_MOUSE
21	157	23.6	1456	1	MANR_HUMAN
22	154.5	23.2	321	1	FCE2_HUMAN
23	153.5	23.0	304	1	MMGL_MOUSE
24	153	23.0	548	1	KOCR_MOUSE
25	150.5	22.6	311	1	LECI_HUMAN
26	149	22.4	207	1	LECH_CHICK
27	143.5	21.5	301	1	LECI_MOUSE
28	141.5	21.2	301	1	LECI_RAT
29	141.5	21.2	331	1	FCE2_MOUSE
30	133.5	20.0	306	1	MMGL_RAT
31	133	20.0	158	1	LECG_TRIST
32	132	19.8	550	1	KOCR_RAT
33	129.5	19.4	283	1	LECH_MOUSE

34	126	18.9	374	1	PSPD_MOUSE
35	125	18.8	374	1	PSPD_RAT
36	123.5	18.5	244	1	MABC_MOUSE
37	122	18.3	162	1	LECG3_MEGMO
38	121	18.2	165	1	LITH_RAT
39	121	18.2	2124	1	PCGA_RAT
40	121	18.2	2132	1	PCGA_MOUSE
41	120.5	18.1	283	1	LECH_RAT
42	120	18.0	2333	1	PCGA_CANFA
43	119.5	17.9	273	1	MT75_MOUSE
44	118.5	17.8	2415	1	PCGA_HUMAN
45	118	17.7	3562	1	PCGV_CHICK

ALIGNMENTS

RESULT ID	1	LEM1_HUMAN	STANDARD:	PRT:	372 AA.
AC	PIA151	PIA5023			
DT	01-JAN-1990	(Rel. 13, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte surface antigen Leu-8) (TQ1) (gp90-MEL)				
DE	(Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).				
GN	SELL OR LYAM1 OR LNHR.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89315837; PubMed=2664786;				
RA	Siegelman M.H., Weissman I.L.;				
RT	"Human homologue of mouse lymph node homing receptor: evolutionary conservation at tandem cell interaction domains."				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:5562-5566(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Tonsil;				
RX	MEDLINE=89310350; PubMed=2473156;				
RA	Tedder T.F., Isaacs C.M., Ernst T.J., Demetri G.D., Adler D.A.,				
RT	Distech C.M.;				
RT	"Isolation and chromosomal localization of cDNAs encoding a novel human lymphocyte cell surface molecule, LAM-1. Homology with the mouse lymphocyte homing receptor and other human adhesion proteins."				
RL	J. Exp. Med. 170:123-133(1989).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90044046; PubMed=2509939;				
RA	Camerlini D., James S.P., Stamenkovic I., Seed B.;				
RT	"Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor."				
RL	Nature 342:78-82(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89308881; PubMed=2663882;				
RA	Bowen B.R., Nguyen T., Lasky L.A.;				
RT	"Characterization of a human homologue of the murine peripheral lymph node homing receptor."				
RL	J. Cell Biol. 109:421-427(1989).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90243637; PubMed=1692315;				
RA	Ord D.C., Ernst T.J., Zhou L.J., Rambaldi A., Spertini O.,				
RT	Griffin J., Tedder T.F.;				
RT	"Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils."				
RL	J. Biol. Chem. 265:7760-7767(1990).				
RN	[6]				
RP	3D-STRUCTURE MODELING.				

```

RX MEDLINE=96074584; PubMed=7488174;
RA Bajorath J., Aruffo A.;
RT "A template for generation and comparison of three-dimensional
RT selectin models.";
RL Biochem. Biophys. Res. Commun. 216:1018-1023(1995).
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHT (SCR) DOMAINS.
CC DATABASE: NAME=PROM; NOTE=CD guide CD62L entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd62l.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25280; AAC63053.1; -
DR EMBL: X16150; CAA34275.1; -
DR EMBL: X17519; CAB43536.1; -
DR EMBL: X17519; CAB43537.1; ALT_SEQ.
DR EMBL: X16070; CAA34203.1; ALT_SEQ.
DR EMBL: M32414; AAB60700.1; -
DR EMBL: M32406; AAB60700.1; JOINED.
DR EMBL: M32407; AAB60700.1; JOINED.
DR EMBL: M32408; AAB60700.1; JOINED.
DR EMBL: M32409; AAB60700.1; JOINED.
DR EMBL: M32410; AAB60700.1; JOINED.
DR EMBL: M32411; AAB60700.1; JOINED.
DR EMBL: M32412; AAB60700.1; JOINED.
DR EMBL: M32413; AAB60700.1; JOINED.
DR PIR: A33912; A33912.
DR PIR: A34015; A34015.
DR PIR: S06798; S06798.
DR PIR: J10104; J10104.
DR PDB: 1KJB; 03-APR-96.
DR GLYCOSULEDB: P14151; -
DR MIM: 153240; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; Lectin_c; 1.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_2; 1.
DR Cell adhesion: Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat; 3D-structure.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRASMEX 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
BY SIMILARITY.

```

```

FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 37 37 D -> Y (IN REF. 4).
FT CONFLICT 178 178 Y -> H (IN REF. 4).
FT CONFLICT 193 193 F -> L (IN REF. 1 AND 4).
FT CONFLICT 213 213 P -> S (IN REF. 3).
FT CONFLICT 214 214 L -> F (IN REF. 4).
FT CONFLICT 218 220 SFS -> NFN (IN REF. 2).
FT CONFLICT 242 242 G -> E (IN REF. 2).
SQ SEQUENCE 372 AA; 42187 MW; 6EA9918ECA2D3643 CRC64;

Query Match 100.0%; Score 666; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,8e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WYHYSEKPMNWRARPCNDYTDVAIONKAIEYLEKTLPFSSSYWIGIRKIGTW 60
DB 39 WYHYSEKPMNWRARPCNDYTDVAIONKAIEYLEKTLPFSSSYWIGIRKIGTW 98
OY 61 TWYGTNKSLLTEAEENWGDGEPNNKKKECVETLYTKRNNDACKWDDACHKIKALC 117
DB 99 TWYGTNKSLLTEAEENWGDGEPNNKKKECVETLYTKRNNDACKWDDACHKIKALC 155

RESULT 2
ID LEM1_PANTR STANDARD; PRT; 372 AA.
AC 095237;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (LECAM1) (CD62L).
GN SELL.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Budhan J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsurushita N.;
RA Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHT (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



```

CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: U73729; AAB18248.1; -.
CC HSSP; P14151; IKUB.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR002396; Selectin.
CC InterPro; IPR000436; Sush1_SCR_CCP.
CC InterPro; IPR001304; lectin_c.
CC Pfam; PF00008; EGF_1.
CC Pfam; PF00059; lectin_c_1.
CC Pfam; PF00084; sush1_2.
CC PRINTS; PR00343; SELECTIN.
CC SMART; SM00032; CCP_2.
CC SMART; SM00034; CLECT_1.
CC SMART; SM00181; EGF_1.
CC PROSITE; PS00022; EGF_1_1.
CC PROSITE; PS01186; EGF_2_1.
CC PROSITE; PS00615; C-TYPE LECTIN_1; 1.
CC PROSITE; PS00615; C-TYPE LECTIN_2; 1.
CC Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sush1; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 372
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFD 57 155
FT DISULFD 128 147
FT DISULFD 160 171
FT DISULFD 165 180
FT DISULFD 182 191
FT DISULFD 197 241
FT DISULFD 227 254
FT DISULFD 259 303
FT DISULFD 289 316
FT CARBOHYD 60 104
FT CARBOHYD 104 104
FT CARBOHYD 116 177
FT CARBOHYD 216 216
FT CARBOHYD 232 232
FT CARBOHYD 246 246
FT CARBOHYD 271 271
SQ SEQUENCE 372 AA; 42188 MM; 6EA991802AD3643 CRC64;

Query Match 100.0%; Score 666; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.8e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

GN SELL.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Creniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsurushita N., to the EMBL/GenBank/DBJ databases.
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: U73729; AAB18247.1; -.
CC HSSP; P14151; IKUB.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR002396; Selectin.
CC InterPro; IPR000436; Sush1_SCR_CCP.
CC InterPro; IPR001304; lectin_c.
CC Pfam; PF00008; EGF_1.
CC Pfam; PF00059; lectin_c_1.
CC Pfam; PF00084; sush1_2.
CC PRINTS; PR00343; SELECTIN.
CC SMART; SM00032; CCP_2.
CC SMART; SM00034; CLECT_1.
CC SMART; SM00181; EGF_1.
CC PROSITE; PS00022; EGF_1_1.
CC PROSITE; PS01186; EGF_2_1.
CC PROSITE; PS00615; C-TYPE LECTIN_1; 1.
CC PROSITE; PS00615; C-TYPE LECTIN_2; 1.
CC Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sush1; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFD 57 155
FT DISULFD 128 147
FT DISULFD 160 171
FT DISULFD 165 180
FT DISULFD 182 191
FT DISULFD 197 241
FT DISULFD 227 254
FT DISULFD 259 303
FT DISULFD 289 316
FT CARBOHYD 60 104
FT CARBOHYD 104 104
FT CARBOHYD 116 177
FT CARBOHYD 216 216
FT CARBOHYD 232 232
FT CARBOHYD 246 246
FT CARBOHYD 271 271
SQ SEQUENCE 372 AA; 42188 MM; 6EA991802AD3643 CRC64;

```

SO SEQUENCE 372 AA; 42118 MW; 6517DD22213FF15E CRC64;

Query Match
Best Local Similarity 99.4%; Score 662; DB 1; Length 372;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNMOKARFRPCRDNTDVLVAIONKAEIYELEKTLPFPSRSYWGIRKIGIW 60
DB 39 WTYHYSEKPMNMOKARFRPCRDNTDVLVAIONKAEIYELEKTLPFPSRSYWGIRKIGIW 98

OY 61 TWVGTNKSILTEAEENNGDGEPPNNKKKEDCEVEIYIRKRDAGKWNDDACHKIKAAALC 117
DB 99 TWVGTNKSILTEAEENNGDGEPPNNKKKEDCEVEIYIRKRDAGKWNDDACHKIKAAALC 155

RESULT 4
LEML_MACMU STANDARD; PRT; 372 AA.

AC 095198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN SELL.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; NCBI_Taxid=9544;
OX NCBI_Taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsunashita N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U73730; AAB18246.1; -
CC HSSP: P14151; 1KDB.
DR InterPro: IPR000361; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; Lectin_C_1.
DR Pfam: PF00084; Sush1_2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP_2.
DR SMART: SM00034; CLECT_1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin; Selectin; Signal; Sush1; Repeat.

FT SIGNAL 1 28 BY SIMILARITY.
FT PROPEP 29 38 BY SIMILARITY.
FT CHAIN 39 372 L-SELECTIN.
FT DOMAIN 39 332 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 333 355 POTENTIAL.
FT DOMAIN 356 372 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 156 192 EGF-LIKE.
FT DOMAIN 196 255 SUSHI 1.
FT DOMAIN 258 317 SUSHI 2.
FT DISULFID 37 135 BY SIMILARITY.
FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT CARBOHYD 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 372 AA; 42109 MW; 08BD7DD5AC549D6D CRC64;

Query Match
Best Local Similarity 94.7%; Score 631; DB 1; Length 372;
Matches 110; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNMOKARFRPCRDNTDVLVAIONKAEIYELEKTLPFPSRSYWGIRKIGIW 60
DB 39 WTYHYSEKPMNMOKARFRPCRDNTDVLVAIONKAEIYELEKTLPFPSRSYWGIRKIGIW 98

OY 61 TWVGTNKSILTEAEENNGDGEPPNNKKKEDCEVEIYIRKRDAGKWNDDACHKIKAAALC 117
DB 99 TWVGTNKSILTEAEENNGDGEPPNNKKKEDCEVEIYIRKRDAGKWNDDACHKIKAAALC 155

RESULT 5
LEML_PAPHA STANDARD; PRT; 372 AA.

AC 028768;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN SELL.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; NCBI_Taxid=9557;
OX NCBI_Taxid=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-97128794; PubMed-8973334;
RA Tsunashita N., Fu H., Berg E.L.;
RT PCR cloning of the cDNA encoding baboon L-selectin.";
RL Gene 181:219-220(1996).
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U52074; AAB40903.1; -
DR HSSP: P14151; 1KJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C_Type_Lectin_1; 1.
DR PROSITE: PS00041; C_Type_Lectin_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28 BY SIMILARITY.
FT PROPEP 29 38 BY SIMILARITY.
FT CHAIN 39 372 L-SELECTIN.
FT DOMAIN 39 332 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 333 355 POTENTIAL.
FT DOMAIN 356 372 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 156 192 EGF-LIKE.
FT DOMAIN 196 255 SUSHT 1.
FT DOMAIN 258 317 SUSHT 2.
FT DISULFID 57 155 BY SIMILARITY.
FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 42091 MW; 64E7BDD5AC549D69 CRC64;

Query Match 94.7%; Score 631; DB 1; Length 372;
Best Local Similarity 94.0%; Pred. No. 1.le-57;
Matches 110; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 WTYYHSEKPMNMQARARPCRDNYTDLVAIONAKAEYELEKTPFSRSYWGIRKIGIM 60
DB 39 WTYYHSENPMMQKARRCRENTDVAIONAKAEYELEKTPFSRSYWGIRKIGIM 98
OY 61 TWVGTNKSLTGEEAENWGDEPNKKRKNKEDCVETIRKNDACKWMDACKIKLALC 117
DB 99 TWVGTNKSLTGEEAENWGDEPNKKRKNKEDCVETIRKNDACKWMDACKIKRKAALC 155

RESULT 6
LEMI_BOVIN
ID LEMI_BOVIN STANDARD; PRT; 370 AA.
AC P96131;

DT 01-FEB-1996 (Rel. 33; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 01-FEB-1996 (Rel. 33; Last annotation update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (LECAM1) (CD62L).
DE SEL.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92164727; PubMed=1371468;
RA Walcheck B., White M., Kurk S., Kishimoto T.K., Jutla M.A.;
RT "Characterization of the bovine peripheral lymph node homing
RT receptor: a lectin cell adhesion molecule (LECAM).";
RL Eur. J. Immunol. 22:469-476(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94055053; PubMed=7694420;
RA Bosworth B.T., Dowbenko D., Shuster D.E., Harp J.A.;
RT "Bovine L-selectin: a peripheral lymphocyte homing receptor";
RL Vet. Immunol. Immunopathol. 37:201-215(1993).
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTIN/LECAMs.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHT (SCR) DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X62882; CAA44676.1; -
DR HSSP: P14151; 1KJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c; 1.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C_Type_Lectin_1; 1.
DR PROSITE: PS00041; C_Type_Lectin_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28 BY SIMILARITY.
FT PROPEP 29 38 BY SIMILARITY.
FT CHAIN 39 370 L-SELECTIN.
FT DOMAIN 39 333 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 334 354 POTENTIAL.
FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 156 192 EGF-LIKE.
FT DOMAIN 196 255 SUSHT 1.
FT DOMAIN 258 317 SUSHT 2.
FT DISULFID 57 155 BY SIMILARITY.
FT DISULFID 128 147 BY SIMILARITY.

[illegible]

CC	use by non-profit institutions as long as its content is in no way modified and this statement is not removed.
CC	modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/isb-cc) or send an email to license@isb.ch .
CC	or send an email to license@isb.ch .
DR	EMBL; M87861; AAA40008.1; -;
DR	M272332; AAA37712.1; -;
DR	PTR; A42755; A42755.
DR	HSSP; P16109; IFSB.
DR	MGD; MG1:98280; Selp.
DR	InterPro: IPRO000561; EGF_1like.
DR	InterPro: IPRO000742; EGF_2.
DR	InterPro: IPRO003396; Selectin.
DR	InterPro: IPRO004366; Sush1_SCC_CCP.
DR	InterPro: IPRO01304; Lectin_c.
DR	Pfam; PF00008; EGF_1.
DR	Pfam; PF00059; lectin_c; 1.
DR	Pfam; PF00084; sush1; 8.
DR	PRINTS; PR00343; SELECTIN.
DR	SMART; SM00032; CCP; 8.
DR	SMART; SM00034; CLECT; 1.
DR	SMART; SM00181; EGF; 1.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin; Selectin; Signal; Sush; Repeat; Lipoprotein; Palmitate.
FT	SIGNAL 1 41 POTENTIAL.
FT	CHAIN 42 768 P-SELECTIN.
FT	DOMAIN 42 709 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 710 733 POTENTIAL.
FT	DOMAIN 734 768 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 58 158 C-TYPE LECTIN (SHORT FORM).
FT	DOMAIN 159 195 EGF-LIKE.
FT	DOMAIN 199 258 SUSHI 1.
FT	DOMAIN 261 320 SUSHI 2.
FT	DOMAIN 323 382 SUSHI 3.
FT	DOMAIN 385 444 SUSHI 4.
FT	DOMAIN 447 506 SUSHI 5.
FT	DOMAIN 509 568 SUSHI 6.
FT	DOMAIN 579 638 SUSHI 7.
FT	DOMAIN 641 700 SUSHI 8.
FT	DISULFID 60 158 BY SIMILARITY.
FT	DISULFID 131 150 BY SIMILARITY.
FT	DISULFID 163 174 BY SIMILARITY.
FT	DISULFID 168 183 BY SIMILARITY.
FT	DISULFID 185 194 BY SIMILARITY.
FT	DISULFID 200 244 BY SIMILARITY.
FT	DISULFID 230 257 BY SIMILARITY.
FT	DISULFID 262 306 BY SIMILARITY.
FT	DISULFID 292 319 BY SIMILARITY.
FT	DISULFID 324 368 BY SIMILARITY.
FT	DISULFID 354 381 BY SIMILARITY.
FT	DISULFID 386 430 BY SIMILARITY.
FT	DISULFID 416 443 BY SIMILARITY.
FT	DISULFID 448 492 BY SIMILARITY.
FT	DISULFID 478 505 BY SIMILARITY.
FT	DISULFID 510 554 BY SIMILARITY.
FT	DISULFID 540 567 BY SIMILARITY.
FT	DISULFID 580 624 BY SIMILARITY.
FT	DISULFID 610 637 BY SIMILARITY.
FT	DISULFID 642 686 BY SIMILARITY.
FT	DISULFID 672 699 BY SIMILARITY.
FT	CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 654 654 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 661 661 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	LIPID 745 745 PALMITATE (BY SIMILARITY).
FT	SITE 756 759 ENDOCYTOSIS SIGNAL (PROBABLE).
FT	CONFLICT 724 724 A -> E (IN REF. 2).
QO	SEQUENCE 768 AA; 83098 MW; ES17307402f66e68 CRC64;

```

Query Match          68.3%:  Score 455;  DB 1:  Length 768;
Best Local Similarity 67.5%:  Pred. No.3,4e-39;
Matches 79;  Conservative 10;  Mismatches 28;  Indels 0;  Gaps 0;

Qy      1  WTGHSEKPMNQARRPCRDNYDTDLVAIQNAELEYLEKTLTPEFSRYWIGIRKIGIW 60
Db      42  WTGYNSTKAYSNMNSRVCFRCRHFTDLVAIQNNKEIHLNDIVPFNSYWIGIRKINMK 101
Qy      61  TWVGYNKSTTEEAENWGDGEPNNKKNKEDCEYITIKRKNDAGKWNDACHKILKALC 117
Db      102  TWVGYNKTLTEEAENWADNEPNKKNKNDCEYITIKSNSAPGKNWDECFRRKALC 158

RESULT 11
LEMS3_BOVIN
ID      LEMS3_BOVIN      STANDARD:  PRT;  646 AA.
AC      P42201;
DR      01-NOV-1995 (Rel. 32, Created)
DR      01-NOV-1995 (Rel. 32, Last sequence update)
DR      15-JUL-1999 (Rel. 38, Last annotation update)
DE      P-selectin precursor (granule membrane protein 140) (GMP-140) (PAOGEW)
DE      (CD62p) (leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN      SELP.
OS      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Bovinae; Bos.
OX      NCBI_TaxId=9913;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Capillary endothelium;
RA      MEDLINE=93249394; PubMed=7683458;
RA      Strubel N.A., Nguyen M., Kansas G.S., Tedder T.F., Bischoff J.;
RT      "Isolation and characterization of a bovine cDNA encoding a
        functional homolog of human P-selectin.";
RL      Biochem. Biophys. Res. Commun. 192:338-344(1993).
CC      -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS
        TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
        INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
        LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
        AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
        ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
        THE CELL SURFACE.
CC      -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC      -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      -1- SIMILARITY: CONTAINS 6 SUSHI (SCD) DOMAINS; BOVINE P-LECTIN LACKS
        THE HUMAN SUSHI-3, -4 AND -7 EQUIVALENTS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: L12041; AAA30743.1;  -.
DR      HSPB; P16109; IESB.
DR      InterPro: IPR000561; EGF-like.
DR      InterPro: IPR002396; Selectin.
DR      InterPro: IPR00436; Sushi_SCR_CCP.
DR      InterPro: IPR01304; lectin_c.
DR      Pfam: PF00008; EGF_1.
DR      Pfam: PF00059; lectin_c_1.
DR      Pfam: PF00084; sushi_6.
DR      PRINTS: PR00343; SELECTIN.
DR      SMART: SM00032; CCP; 6.
DR      SMART: SM00034; CLECT; 1.

```


DR SMART: SM00181; EGF_1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN_2; 1.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sushi; Repeat.
 FT SIGNAL 1 41
 FT CHAIN 1 41
 FT DOMAIN 42 646
 FT TRANSMEM 588 611
 FT DOMAIN 612 646
 FT DOMAIN 58 158
 FT DOMAIN 159 195
 FT DOMAIN 199 258
 FT DOMAIN 261 320
 FT DOMAIN 323 382
 FT DOMAIN 385 444
 FT DOMAIN 457 516
 FT DOMAIN 519 578
 FT DISULFID 60 158
 FT DISULFID 131 150
 FT DISULFID 163 174
 FT DISULFID 168 183
 FT DISULFID 185 194
 FT DISULFID 200 244
 FT DISULFID 230 257
 FT DISULFID 262 306
 FT DISULFID 292 319
 FT DISULFID 324 368
 FT DISULFID 354 381
 FT DISULFID 386 430
 FT DISULFID 416 443
 FT DISULFID 458 502
 FT DISULFID 488 515
 FT DISULFID 520 564
 FT DISULFID 550 577
 FT CARBOHYD 48 48
 FT CARBOHYD 54 54
 FT CARBOHYD 80 80
 FT CARBOHYD 180 180
 FT CARBOHYD 212 212
 FT CARBOHYD 219 219
 FT CARBOHYD 336 336
 FT CARBOHYD 481 481
 FT CARBOHYD 532 532
 FT CARBOHYD 539 539
 FT CARBOHYD 557 557
 FT SITE 634 637
 SQ SEQUENCE 646 AA: 71229 MW: 573912A4627A6ACA CRC64;

Query Match 67.7%; Score 451; DB 1; Length 646;
 Best local similarity 66.7%; Pred. No. 7.3e-39;
 Matches 78; Conservative 10; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTHYSEKPMNORARFCEDNTDVAIONKAEILEKTLPESSSYWIGIRKIGIM 60
 DB 42 WTHYSEKPMNORARFCEDNTDVAIONKAEILEKTLPESSSYWIGIRKINNM 101
 OY 61 TWVGTKKTLTEEAENMGDGEPPNKKKEDCVETIKRNNDAGWMDACHKLKAALC 117
 DB 102 TWVGTKKTLTEEAENMGDGEPPNKKKEDCVETIKRNNDAGWMDACHKLKAALC 158

RESULT 12
 LEM2_BOVIN ID LEM2_BOVIN STANDARD; PRT; 485 AA.
 AC P98107:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)

DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
 DE (CD62e).
 GN SELE.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Adrenal gland;
 RX MEDLINE=93382537; PubMed=7690465;
 RA Nguyen M., Strubel N.A., Bischoff J.;
 RT "A role for sialyl Lewis x/A glycoconjugates in capillary morphogenesis";
 RL Nature 365:267-269(1993).
 CC -!- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF POLYGLYCOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF GLYCOPOLIDS).
 CC -!- FUNCTION: INVOLVED IN CAPILLARY MORPHOGENESIS BY BINDING TO A BOVINE CAPILLARY ENDOTHELIAL (BCE) CELL SIALYL LEWIS-X AND/OR SIALYL LEWIS A-CONTAINING LIGAND (PROBABLY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: TO OTHER SELECTINS/LECTINS.
 CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAINS; BOVINE E-LECTIN LACKS THE HUMAN SUSHI-4 AND 5 EQUIVALENTS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L12039; AAA02991.1; -
 CC HSP: P16581; IEST.
 CC DR InterPro: IPR000561; EGF-like.
 CC DR InterPro: IPR000742; EGF_2.
 CC DR InterPro: IPR002396; Selectin.
 CC DR InterPro: IPR000436; Sushi_SCR_CCP.
 CC DR InterPro: IPR001304; Lectin_c.
 CC Pfam: PF00008; EGF_1.
 CC Pfam: PF00059; lectin_c; 1.
 CC Pfam: PF00084; sushi; 4.
 CC PRINTS: PR00343; SELECTIN.
 CC SMART: SM00032; CCP; 4.
 CC SMART: SM00034; CLECT; 1.
 CC SMART: SM00181; EGF_1; 1.
 CC PROSITE: PS00022; EGF_1; 1.
 CC PROSITE: PS01186; EGF_2; 1.
 CC PROSITE: PS00615; C-TYPE LECTIN_1; 1.
 CC PROSITE: PS00615; C-TYPE LECTIN_2; 1.
 CC Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sushi; Repeat.
 FT SIGNAL 1 22
 FT CHAIN 23 485
 FT DOMAIN 23 430
 FT TRANSMEM 431 453
 FT DOMAIN 454 485
 FT DOMAIN 39 139
 FT DOMAIN 140 176
 FT DOMAIN 180 238
 FT DOMAIN 241 300
 FT DOMAIN 303 363
 FT DOMAIN 366 422
 FT DOMAIN 41 139
 FT DISULFID 112 131
 FT DISULFID 144 155

RP STRUCTURE BY NMR OF 160-199.
RX MEDLINE=97057176; PubMed=8901515;
RA Freedman S.D., Sanford D.G., Bachovchin W.W., Furie B.C., Baleja J.D.,
RT Furie B.;
RT "Structure and function of the epidermal growth factor domain of P-
RT selectin.";

```

RL Biochemistry 35:13733-13744(1996).
RN [4]
RX 3D-STRUCTURE MODELING OF 42-161.
RX MEDLINE=94093388; PubMed=7505680;
RA Bajorath J., Stenkamp R., Aruffo A.;
RT "Knowledge-based model building of proteins: concepts and examples.";
RL Protein Sci. 2:1798-1810(1993).
RN [5]
RX VARIANTS ASN-331; ASP-603; VAL-640 AND PRO-756.
RX MEDLINE=98334547; PubMed=9668170;
RA Hermann S.M., Ricard S., Nicoud V., Mallet C., Evans A.,
RA Ruidavets J.B., Arveiler D., Luc G., Cambien F.;
RT "The P-selectin gene is highly polymorphic: reduced frequency of the
RT pro175 allele carriers in patients with myocardial infarction.";
RL Hum. Mol. Genet. 7:1277-1284(1998).
CC -1- FUNCTION: CA12+3-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC LEUKOCYTES. THE LIGAND RECOGNIZED IS STALY1-LEWIS X.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
CC THE CELL SURFACE.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 9 SUSHI (SCR) DOMAINS.
CC -1- DATABASE: NAME=PROX, NOTE=CD guide CD62p entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prox/cd/cd62p.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60234; AAA35910.1; -.
DR EMBL: M60217; AAA35910.1; JOINED.
DR EMBL: M60218; AAA35910.1; JOINED.
DR EMBL: M60219; AAA35910.1; JOINED.
DR EMBL: M60222; AAA35910.1; JOINED.
DR EMBL: M60223; AAA35910.1; JOINED.
DR EMBL: M60224; AAA35910.1; JOINED.
DR EMBL: M60225; AAA35910.1; JOINED.
DR EMBL: M60226; AAA35910.1; JOINED.
DR EMBL: M60227; AAA35910.1; JOINED.
DR EMBL: M60228; AAA35910.1; JOINED.
DR EMBL: M60229; AAA35910.1; JOINED.
DR EMBL: M60231; AAA35910.1; JOINED.
DR EMBL: M60232; AAA35910.1; JOINED.
DR EMBL: M60233; AAA35910.1; JOINED.
DR EMBL: M25322; AAA35911.1; JOINED.
DR PIR: A30359; A30359.
DR PDB: 1FSB: 01-APR-97.
DR PDB: 1KJD: 03-APR-96.
DR MIM: 173610; -.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushi; 9.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 9.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE LECTIN1; 1.

```


OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX	NCBI_TaxID=9986;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Heart;
RX	MEDLINE=92189729; PubMed=1372169;
RA	Lairigan J.D., Tseng T.C., Rumberger J.M., Burns D.K.;
RT	"Characterization of cDNA and genomic sequences encoding rabbit
RT	ELAM-1: conservation of structure and functional interactions with
RT	leukocytes.";
RL	DNA Cell Biol. 11:149-162(1992).
CC	-1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC	MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC	ELAM-1 IS SIMILY-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC	POLYACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC	GLYCOPIDS).
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- INDUCTION: BY CYTOKINES.
CC	-1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC	-1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed, usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M91004; AAA31243.1; -;
DR	EMBL; M91005; AAA31244.1; -;
DR	HSSP; P16583; IKTA.
DR	InterPro; IPR000561; EGF-like.
DR	InterPro; IPR002396; Selectin.
DR	InterPro; IPR000436; Sushi_SCR_CCP.
DR	InterPro; IPR001304; lectin_c.
DR	Pfam; PF00008; EGF_1.
DR	Pfam; PF00059; lectin_c_1.
DR	Pfam; PF00084; sushi_5.
DR	PRINTS; PR00343; SELECTIN.
DR	SMART; SM00032; CCP; 5.
DR	SMART; SM00034; CLECT; 1.
DR	SMART; SM00181; EGF_1.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR	PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW	Selectin; Signal; Sushi; Repeat.
FT	SIGNAL 1 23
FT	CHAIN 1 551
FT	DOMAIN 24 495 E-SELECTIN.
FT	TRANSMEM 496 517 EXTRACELLULAR (POTENTIAL).
FT	DOMAIN 518 551 POTENTIAL.
FT	DOMAIN 40 140 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 141 177 C-TYPE LECTIN (SHORT FORM).
FT	DOMAIN 181 240 EGF-LIKE.
FT	DOMAIN 243 302 SUSHI 1.
FT	DOMAIN 305 365 SUSHI 2.
FT	DOMAIN 368 428 SUSHI 3.
FT	DOMAIN 431 487 SUSHI 4.
FT	DISULFID 431 487 SUSHI 5.
FT	DISULFID 42 140 BY SIMILARITY.
FT	DISULFID 113 132 BY SIMILARITY.
FT	DISULFID 145 156 BY SIMILARITY.
FT	DISULFID 150 165 BY SIMILARITY.
FT	DISULFID 167 176 BY SIMILARITY.
FT	DISULFID 182 226 BY SIMILARITY.
FT	DISULFID 212 239 BY SIMILARITY.
FT	DISULFID 244 288 BY SIMILARITY.
FT	DISULFID 274 301 BY SIMILARITY.


```
FT CONFLICT 327 327 K -> N (IN REF. 2).
FT CONFLICT 363 363 V -> A (IN REF. 2).
FT CONFLICT 384 384 V -> M (IN REF. 2).
FT CONFLICT 461 484 KFPSSSESLQPGSGSTOMEDDLI ->
FT SEQUENCE 484 AA: 52567 MM: AEF74FE25C1ED013 CRC64;
SQ

Query Match 63.88; Score 425; DB 1; Length 484;
Best Local Similarity 63.28; Pred. No. 2.5e-36;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WYHSEKPMWMMORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYWMIGIRKIGTW 60
OY 23 WSYSTSTETMTFDASATCCOQRTYHLVAIONHAEIYEIYNSTFNYSASTYWGIRKINGTW 82
OY 61 TWVGTNKSITTEAEVNMWGDGEPNNKKNKEDCVEIYIKRNKADGKWDNDCKHKLKALC 117
OY 83 TWIGTKKALTEPATWMAWGPENNNKOSNEDCVEIYIKRDKSGKNWDERCSKKKLLALC 139
OY

RESULT 17
LEM2_HUMAN STANDARD; PRT; 610 AA.
AC P16581; P16111;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE E-selectin precursor (endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE OR ELAM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-90175359; PubMed-1689848;
RA Hession C., Osborn L., Goff D., Chi-Rosso G., Vassallo C.,
RA Parek M., Pltack C., Tizard R., Goetz S., McCarthy K., Hopple S.,
RA Lobb R.;
RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning
RT and functional interactions."
RL Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE-89162047; PubMed-2466335;
RA Bevilacqua M.P., Stengelin S., Gimbrome M.A. Jr., Seed B.;
RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for
RT neutrophils related to complement regulatory proteins and lectins."
RL Science 243:1160-1165(1989).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE-9115870; PubMed-1703529;
RA Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T.,
RA Gimbrome M.A. Jr., Bevilacqua M.P.;
RT "Structure and chromosomal location of the gene for endothelial-
RT leukocyte adhesion molecule 1."
RL J. Biol. Chem. 266:2466-2473(1991).
RN 14
RP LIGAND.
RX MEDLINE-91068005; PubMed-1701274;
RA Phillips M.L., Nudelmaier E., Gaeta F.C., Perez M., Singhal A.K.,
RA Hakomori S., Paulson J.C.;
RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate
RT ligand, sialyl-Lex."
RL Science 250:1130-1132(1990).
RN 15
RP 3D-STRUCTURE MODELLING OF LECTIN DOMAIN.
RX MEDLINE-93202275; PubMed-7681016;
RA Mills A.;
RT "Modelling the carbohydrate recognition domain of human E-selectin.";
```

```
RL FEBS Lett. 319:5-11(1993).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.
RX MEDLINE-94150646; PubMed-7509040;
RA Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S.,
RA Huang K.-S., Presky D.H., Familletti P.C., Wollitzky B.A., Burns D.K.;
RT "Insight into E-selectin/ligand interaction from the crystal
RT structure and mutagenesis of the Lec/EGF domains."
RL Nature 367:532-538(1994).
RN [7]
RP VARIANT ARG-149.
RX MEDLINE-95179107; PubMed-7533025;
RA Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S.,
RA Schulte K.L., Glaser C., Rohde K., Baumann G., Speer A.;
RT "E-selectin polymorphism and atherosclerosis: an association study."
RL Hum. Mol. Genet. 3:1935-1937(1994).
RN [8]
RP VARIANT ARG-149.
RX MEDLINE-99134508; PubMed-9933738;
RA Ye S.Q., Usher D., Virgil D., Zhang L.Q., Yochim S.E., Gupta R.;
RT "A PstI polymorphism detects the mutation of serine-128 to arginine in
RT CD 62e gene - a risk factor for coronary artery disease."
RL J. Biomed. Sci. 6:18-21(1999).
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1-3)FUCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A
CC HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY
CC HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH
CC ANGIOGRAPHICALLY PROVEN SEVERE ATHEROSCLEROSIS COMPARED WITH AN
CC UNSELECTED POPULATION (SER-149).
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
CC -1- DATABASE: NAME=PROW, NOTE=CD guide CD62E entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62e.htm".
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M30640; AAA52377.1; -;
DR EMBL; M61893; AAA52375.1; -;
DR EMBL; M61895; AAA52375.1; JOINED.
DR EMBL; M61887; AAA52375.1; JOINED.
DR EMBL; M61888; AAA52375.1; JOINED.
DR EMBL; M61890; AAA52375.1; JOINED.
DR EMBL; M61891; AAA52375.1; JOINED.
DR EMBL; M61892; AAA52375.1; JOINED.
DR EMBL; M24736; AAA52376.1; -;
DR PIR; A32606; A32606.
DR PIR; A35046; A35046.
DR PIR; A38615; A38615.
DR PDB; 1EST; 31-AUG-94.
DR PDB; 1KJA; 03-APR-96.
DR MIM; 131210; -;
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushl_SCR_CCP.
DR InterPro: IPR001304; Lectin_C.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00059; Lectin_C_1.
DR Pfam; PF00084; sushl_6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 6.
```

```

DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat; Polymorphism; 3d-structure.
FT SIGNAL 1 21
FT CHAIN 22 610
FT DOMAIN 22 556
FT TRANSMEM 557 578
FT DOMAIN 579 610
FT DOMAIN 38 138
FT DOMAIN 139 175
FT DOMAIN 179 238
FT DOMAIN 241 300
FT DOMAIN 303 363
FT DOMAIN 366 426
FT DOMAIN 429 489
FT DOMAIN 492 548
FT DISULFID 40 138
FT DISULFID 111 130
FT DISULFID 143 154
FT DISULFID 148 163
FT DISULFID 165 174
FT DISULFID 180 224
FT DISULFID 210 237
FT DISULFID 242 286
FT DISULFID 272 299
FT DISULFID 304 349
FT DISULFID 335 362
FT DISULFID 367 412
FT DISULFID 398 425
FT DISULFID 430 475
FT DISULFID 461 486
FT DISULFID 493 534
FT DISULFID 520 547
FT CARBOHYD 25 25
FT CARBOHYD 145 145
FT CARBOHYD 160 160
FT CARBOHYD 179 179
FT CARBOHYD 199 199
FT CARBOHYD 203 203
FT CARBOHYD 265 265
FT CARBOHYD 312 312
FT CARBOHYD 332 332
FT CARBOHYD 503 503
FT CARBOHYD 527 527
FT VARIANT 130 130
FT VARIANT 149 149
FT VARIANT 149 149
FT VARIANT 295 295
FT VARIANT 421 421
FT VARIANT 468 468
FT VARIANT 575 575
SQ SEQUENCE 610 AA; 66655 MW; 7D43E3C0D1229229 CRC64;

```

Query Match 62.3%; Score 415; DB 1; Length 610;
 Best Local Similarity 60.7%; Pred. No. 3-5e-35;
 Matches 71; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

```

QY 1 WHYHSEKPMNORARCRDNTDVAIONAEIYLEKTLPEFSRYWIGIRKIGTW 60
DB 22 WSYNTSTAMFYDEASAYCOQRYTHVAIONKEEIEYLNSILSYSPYWIGIRKNNW 81

```

```

QY 61 TWGTSKSLTEAEENMGDEPNKKNKEDCEYIYIKRNDAGKWNPDACHKTLKAALC 117
DB 82 WMTQTRPLTBEAKNMAPGPENNRKDEDCVEYIYIKRNDYGMANDRCKKLLALC 138

```

RESULT 18
 LEM2_RAT
 ID LEM2_RAT STANDARD; PRT; 549 AA.

AC P98105;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
 DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
 DE (CD62E)
 GN SELE OR ELAM-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
 RL Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC ELAM-1 IS STAYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOLIPIDS).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; I25527; AAA4113.1; -;
 DR HSSP; P16581; IKTA.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR InterPro; IPR001304; Lectin_c.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00059; lectin_c; 1.
 DR Pfam; PF00084; sushi; 5.
 DR SMART; SM00032; CCP; 5.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sushi; Repeat.
 FT SIGNAL 1 21
 FT CHAIN 22 549
 FT DOMAIN 22 494
 FT TRANSMEM 495 516
 FT DOMAIN 517 549
 FT DOMAIN 38 138
 FT DOMAIN 139 175
 FT DOMAIN 179 239
 FT DOMAIN 242 301
 FT DOMAIN 304 364
 FT DOMAIN 367 427
 FT DOMAIN 430 486
 FT DOMAIN 430 486
 FT SUSHI 5.

```

FT DISULFID 40 138 BY SIMILARITY.
FT DISULFID 111 130 BY SIMILARITY.
FT DISULFID 143 154 BY SIMILARITY.
FT DISULFID 148 163 BY SIMILARITY.
FT DISULFID 165 174 BY SIMILARITY.
FT DISULFID 180 225 BY SIMILARITY.
FT DISULFID 210 238 BY SIMILARITY.
FT DISULFID 243 287 BY SIMILARITY.
FT DISULFID 273 300 BY SIMILARITY.
FT DISULFID 305 350 BY SIMILARITY.
FT DISULFID 336 363 BY SIMILARITY.
FT DISULFID 368 413 BY SIMILARITY.
FT DISULFID 399 426 BY SIMILARITY.
FT DISULFID 431 472 BY SIMILARITY.
FT DISULFID 458 485 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 549 AA; 60079 MW; 85CECDDB7B0144C8 CRC64;

```

```

Query Match 61.3%; Score 408; DB 1; Length 549;
Best Local Similarity 60.7%; Pred. No. 1,6e-34;
Matches 71; Conservative 14; Mismatches 32; Indels 0; Gaps 0;

```

```

OY 1 WYHYHSEKMNQRRARFRDNTDLVATQNAEITYLEKTLPEFSNSYIWIGIRKIGTW 60
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 22 WYNNASSELMTYDEASAYCQDRYTHLVAIONKEEINYLSLRLSPSYWIGIRKNVNW 81
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 61 TWVHTKSLTEEAENMGDEGPNKKKNEPCVEITYIKRNKDGKWNDDACHKIKLAALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 82 INVGTOKPLTEERKKNWAPGEPNNKORNECDVEITYIORPKDSGMWDERCKKKLALC 138
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 19
LEM2_CANFA STANDARD: PRT: 611 AA.
ID LEM2_CANFA
AC P33730.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Jugular vein;
RA Manning A.M., Lane C.L., Auchampach J.A., Kukiela G.L.,
RA Rosenbloom C.L., Anderson D.C.;
RL Submitted (XXY-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC ELAM-1 IS SIMILY-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

CC EMBL: I23087; AAA30843.1; -.
CC HSSP: P16581; IESL.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR002396; Selectin.
CC InterPro: IPR000436; Sush1_SCR_CCP.
CC InterPro: IPR001304; Lectin_C.
CC Pfam: PF00008; EGF_1.
CC Pfam: PF00059; lectin_c; 1.
CC Pfam: PF00084; sush1; 6.
CC PRINTS: PR00343; SELECTIN.
CC SMART: SM00032; CCP; 6.
CC SMART: SM00034; CLECT; 1.
CC SMART: SM00181; EGF; 2.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
CC PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
CC Cell adhesion. Transmembrane; Glycoprotein; EGF-like domain; Lectin;
CC Selectin; Signal; Sush1; Repeat.
KW SIGNAL 1
FT CHAIN 23 611
FT DOMAIN 23 557
FT TRANSMEM 558 579
FT DOMAIN 580 611
FT DOMAIN 39 139
FT DOMAIN 140 176
FT DOMAIN 180 239
FT DOMAIN 242 301
FT DOMAIN 304 364
FT DOMAIN 367 427
FT DOMAIN 430 490
FT DOMAIN 493 549
FT DISULFID 41 139
FT DISULFID 112 131
FT DISULFID 144 155
FT DISULFID 149 164
FT DISULFID 166 175
FT DISULFID 181 225
FT DISULFID 211 238
FT DISULFID 243 287
FT DISULFID 273 300
FT DISULFID 305 350
FT DISULFID 336 363
FT DISULFID 368 413
FT DISULFID 399 426
FT DISULFID 431 476
FT DISULFID 462 489
FT DISULFID 484 521
FT DISULFID 521 548
FT CARBOHYD 26 26
FT CARBOHYD 161 161
FT CARBOHYD 204 266
FT CARBOHYD 266 266
FT CARBOHYD 313 313
FT CARBOHYD 333 333
FT CARBOHYD 528 528
SO SEQUENCE 611 AA; 66314 MW; 35DA9E3DF225E4F6 CRC64;

```

```

Query Match 61.3%; Score 408; DB 1; Length 611;
Best Local Similarity 59.0%; Pred. No. 1,8e-34;
Matches 69; Conservative 18; Mismatches 30; Indels 0; Gaps 0;
OY 1 WYHYHSEKMNQRRARFRDNTDLVATQNAEITYLEKTLPEFSNSYIWIGIRKIGTW 60

```

```

DB 23 WYNNASTEAMTDEASTYQOQRYTHLVAIONOEIKYLSMFTYPTTYWIGIRVNNKW 82
OY 61 TWVGTNKSILTEAEENWGDPEPNKKNECCEVEIYIKRNKDACKWNDDACHKIKALC 117
DB 83 TWIGTOKLPLEEAKWMBEPENNNKONDECCVEIYIKRKDSKWNDCRCKKALC 139

RESULT 20
LEMM_MOUSE STANDARD; PRT; 612 AA.
ID 000690:
AC 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
CN SELE OR ELAM-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=92283265; PubMed=1375914;
RA Becker-Andre M., van Huijsdujnen R.H., Losberger C., Whelan J.,
RA Delamarier J.F.;
RT "Murine endothelial leukocyte-adhesion molecule 1 is a close
RT structural and functional homologue of the human protein.";
RL Eur. J. Biochem. 206:401-411(1992).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=92340571; PubMed=1378846;
RA Weller A., Isenmann S., Vestweber D.;
RT "Cloning of the mouse endothelial selectins. Expression of both E-
RT and P-selectin is inducible by tumor necrosis factor alpha.";
RL J. Biol. Chem. 267:15176-15183(1992).
CC 1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS STAYL-L-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC 1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC 1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC 1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M80778; AAA37547.1;
DR EMBL: M87862; AAA37577.1; ALT_INIT.
DR HSSP: P16581; IKAJ.
DR MGD: MGI:98278; Sele.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sushi_6.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 6.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00022; EGF_1; 1.

```

```

DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE: PS00041; C-TYPE LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 612
FT DOMAIN 22 557
FT TRANSMEM 558 579
FT DOMAIN 580 612
FT DOMAIN 38 138
FT DOMAIN 139 175
FT DOMAIN 179 239
FT DOMAIN 242 301
FT DOMAIN 304 364
FT DOMAIN 367 427
FT DOMAIN 430 490
FT DOMAIN 493 549
FT DISULFID 40 138
FT DISULFID 111 130
FT DISULFID 143 154
FT DISULFID 148 163
FT DISULFID 165 174
FT DISULFID 180 225
FT DISULFID 210 238
FT DISULFID 243 287
FT DISULFID 273 300
FT DISULFID 305 350
FT DISULFID 336 363
FT DISULFID 368 413
FT DISULFID 399 426
FT DISULFID 431 476
FT DISULFID 462 489
FT DISULFID 494 535
FT DISULFID 521 548
FT CARBOHYD 25 25
FT CARBOHYD 145 145
FT CARBOHYD 192 192
FT CARBOHYD 203 203
FT CARBOHYD 266 266
FT CARBOHYD 333 333
FT CARBOHYD 391 391
FT CARBOHYD 504 504
FT CARBOHYD 528 528
SO SEQUENCE 612 AA; 66749 MW; 86F05713F0EC2C3D CRC64;

Query Match 60.8%; Score 405; DB 1; Length 612;
Best Local Similarity 59.8%; Pred. No. 3; 7e-34;
Matches 70; Conservative 14; Mismatches 33; Indels 0; Gaps 0;

OY 1 WYHYSEKPNMQRARPCNDYTDIVATONKAEITYLEKTLPLFSRSYYWIGIRIGGIW 60
DB 22 WYNNASSELMTYDEASAYCQRDTHLVAIONKEEYLSNKHSPSYWIGIRVNNW 81
OY 61 TWVGTNKSILTEAEENWGDPEPNKKNECCEVEIYIKRNKDACKWNDDACHKIKALC 117
DB 82 TWIGTOKLPLEEAKWMBEPENNNKONDECCVEIYIKRKDSKWNDCRCKKALC 139

RESULT 21
MANR_HUMAN STANDARD; PRT; 1456 AA.
ID MANR_HUMAN
AC P22897;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Macrophage mannose receptor precursor (MMR) (CD206 antigen).
GN MRC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;

```


RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP TISSUE=Placenta;
 RC MEDLINE=90324192; PubMed=2373685;
 RA Taylor M.E., Conary J.T., Lemnartz W.R., Stahl P.D., Dickamer K.;
 RT "Primary structure of the mannose receptor contains multiple motifs
 RL resembling carbohydrate-recognition domains";
 RL J. Biol. Chem. 265:12156-12162(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93052405; PubMed=1294118;
 RA Kim S.J., Ruiz N., Bezouska K., Dickamer K.;
 RT "Organization of the gene encoding the human macrophage mannose
 RL receptor (MRC1).";
 RL Genomics 14:721-727(1992).
 RN [3]
 RP STUDIES ON THE BINDING OF INDIVIDUAL LECTIN DOMAINS.
 RX MEDLINE=92112893; PubMed=1730714;
 RA Taylor M.E., Bezouska K., Dickamer K.;
 RT "Contribution to ligand binding by multiple carbohydrate-recognition
 RL domains in the macrophage mannose receptor.";
 RL J. Biol. Chem. 267:1719-1726(1992).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 642-788.
 RX MEDLINE=20347275; PubMed=10779515;
 RA Feinberg H., Park-Snyder S., Kolatkar A.R., Heise C.T., Taylor M.E.,
 RA Weiss W.I.;
 RT "Structure of a C-type carbohydrate recognition domain from the
 RL macrophage mannose receptor";
 RL J. Biol. Chem. 275:21539-21548(2000).
 CC -1 FUNCTION: MEDIATES THE ENDOCYTOSIS OF GLYCOPROTEINS BY
 CC MACROPHAGES, IN SEVERAL RECOGNITION AND UPTAKE PROCESSES.
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 MISCELLANEOUS: CRDS 1-3 HAVE AT MOST VERY WEAK AFFINITY FOR
 CC CARBOHYDRATE. CRD 4 SHOWS THE HIGHEST AFFINITY BINDING AND HAS
 CC MULTISPECIFICITY FOR A VARIETY OF MONOSACCHARIDES. AT LEAST 3 CRDS
 CC (4, 5, AND 7) ARE REQUIRED FOR HIGH AFFINITY BINDING AND
 CC ENDOCYTOSIS OF MULTIVALENT GLYCOCONGUGATES.
 CC -1 SIMILARITY: CONTAINS 8 C-TYPE LECTIN FAMILY DOMAINS.
 CC -1 SIMILARITY: CONTAINS 1 RICHIN B-TYPE LECTIN DOMAIN.
 CC -1 DATABASE: NAME=PROV; NOTE=CD guide CD206 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1044341355_g.htm".
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J05550; AAA59868.1; -;
 DR EMBL: M93221; AAA60389.1; -;
 DR EMBL: M93192; AAA60389.1; JOINED.
 DR EMBL: M93193; AAA60389.1; JOINED.
 DR EMBL: M93194; AAA60389.1; JOINED.
 DR EMBL: M93195; AAA60389.1; JOINED.
 DR EMBL: M93196; AAA60389.1; JOINED.
 DR EMBL: M93197; AAA60389.1; JOINED.
 DR EMBL: M93198; AAA60389.1; JOINED.
 DR EMBL: M93199; AAA60389.1; JOINED.
 DR EMBL: M93200; AAA60389.1; JOINED.
 DR EMBL: M93201; AAA60389.1; JOINED.
 DR EMBL: M93202; AAA60389.1; JOINED.
 DR EMBL: M93203; AAA60389.1; JOINED.
 DR EMBL: M93204; AAA60389.1; JOINED.
 DR EMBL: M93205; AAA60389.1; JOINED.
 DR EMBL: M93206; AAA60389.1; JOINED.
 DR EMBL: M93207; AAA60389.1; JOINED.
 DR EMBL: M93208; AAA60389.1; JOINED.
 DR EMBL: M93209; AAA60389.1; JOINED.
 DR EMBL: M93210; AAA60389.1; JOINED.
 DR EMBL: M93211; AAA60389.1; JOINED.

DR	EMBL:	M93212;	AAA60389.1;	JOINED.
DR	EMBL:	M93213;	AAA60389.1;	JOINED.
DR	EMBL:	M93214;	AAA60389.1;	JOINED.
DR	EMBL:	M93215;	AAA60389.1;	JOINED.
DR	EMBL:	M93216;	AAA60389.1;	JOINED.
DR	EMBL:	M93217;	AAA60389.1;	JOINED.
DR	EMBL:	M93218;	AAA60389.1;	JOINED.
DR	EMBL:	M93219;	AAA60389.1;	JOINED.
DR	EMBL:	M93220;	AAA60389.1;	JOINED.
DR	PfR:	A36563;	A36563.	
DR	PDB:	1EGG;	30-AUG-00.	
DR	PDB:	1EGI;	30-AUG-00.	
DR	MM:	153618;	-	
DR	InterPro:	IPR000562;	FN_Type_II.	
DR	InterPro:	IPR000772;	Ricin_B_lectin.	
DR	InterPro:	IPR01304;	lectin_c.	
DR	Pfam:	PF00040;	fln2; 1.	
DR	Pfam:	PF00059;	lectin_c; 8.	
DR	Pfam:	PF00652;	Ricin_B_lectin; 1.	
DR	PRINTS:	PR00013;	FNTYPEII.	
DR	ProDom:	PD000995;	FN_Type_II; 1.	
DR	SMART:	SM00034;	CLECT; 8.	
DR	SMART:	SM00059;	FN2; 1.	
DR	SMART:	SM00458;	RICIN; 1.	
DR	PROSITE:	PS00615;	C_Type_Lectin_1; 6.	
DR	PROSITE:	PS00641;	C_Type_Lectin_2; 8.	
DR	PROSITE:	PS00023;	FIBRONECTIN_2; 1.	
DR	PROSITE:	PS0231;	RICIN_B_lectin; 1.	
KW	Receptor:	Signal;	Calcium-binding; Transmembrane; Repeat;	
KW	Glycoprotein:	Endocytosis; MacropHage; Lectin; Antigen; 3D-structure.		
KM	SIGNAL	1	18	POTENTIAL.
FT	CHAIN	19	1456	MACROPHAGE MANNOSE RECEPTOR.
FT	DOMAIN	19	1383	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1384	1411	POTENTIAL.
FT	DOMAIN	1412	1456	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	22	142	RICIN B-TYPE LECTIN.
FT	DOMAIN	157	212	FIBRONECTIN TYPE-II.
FT	DOMAIN	216	344	C-TYPE LECTIN 1 (LONG FORM).
FT	DOMAIN	360	490	C-TYPE LECTIN 2 (LONG FORM).
FT	DOMAIN	502	629	C-TYPE LECTIN 3 (LONG FORM).
FT	DOMAIN	644	781	C-TYPE LECTIN 4 (LONG FORM).
FT	DOMAIN	805	926	C-TYPE LECTIN 5 (LONG FORM).
FT	DOMAIN	943	1083	C-TYPE LECTIN 6 (LONG FORM).
FT	DOMAIN	1100	1216	C-TYPE LECTIN 7 (LONG FORM).
FT	DOMAIN	1228	1359	C-TYPE LECTIN 8 (LONG FORM).
FT	DISULFID	646	659	
FT	DISULFID	680	777	
FT	DISULFID	753	769	
FT	CARBOHYD	104	104	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	344	344	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	529	529	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	926	926	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	930	930	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1160	1160	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1205	1205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1311	1311	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	1456 AA;	166011 MW;	26455AF3C576A5E3 CRC64;

Query Match	23.6%	Score 157;	DB 1;	Length 1456;
Best Local Similarity	30.8%	Pred. No. 3,4e-08;		
Matches	37;	Conservative	29;	Mismatches 44; Indels 10; Gaps 4;
QY	1	WTYHYSEKPMNQARRRRCRDNYTDLVAIQNAEILEYEKLT--PFSRSYWGIR-KIG	57	
	:			
DB	810	YQYFSKKEKEMTDNRAAFCKRNFGLDVSIQSESEKFLMKYVNRDAQSAVETGLISLD	869	
QY	58	GWTWVGINKKSLTEAEENNGDGEFPNNKKKRECEVELYIKRNKDACKWMDACHKLKALC	117	
	:			
DB	870	KKFAWMDGSK--VDYVSWATGEPNFANEDKECVTRY---SNSGFMDINICGYAFNAFC	922	
RESULT	22			

FCE2_HUMAN STANDARD; PRT; 321 AA.
 ID FCE2_HUMAN
 AC P06734;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Low affinity immunoglobulin epsilon FC receptor (lymphocyte IgE receptor) (Fc-epsilon-RI) (CD23) (BLAST-2) (immunoglobulin E-binding factor).
 DE FcER2 OR IGE2F.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8711825; PubMed=2949326;
 RA Ikuta K., Takami M., Kim C.W., Honjo T., Miyoshi T., Tagaya Y., Kawabe T., Iodori J.;
 RT "Human lymphocyte Fc receptor for IgE: sequence homology of its cloned cDNA with animal lectins.";
 RT Proc. Natl. Acad. Sci. U.S.A. 84:819-823(1987).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87051737; PubMed=2877743;
 RA Kikuchi H., Inui S., Sato R., Barsumian E.L., Owaki H., Yamasaki K., Kaisho T., Uchiyayashi N., Hardy R.R., Hirano T., Tsunasawa S., Sakiyama F., Suenura M., Kishimoto T.;
 RT "Molecular structure of human lymphocyte receptor for immunoglobulin E.";
 RT Cell 47:657-665(1986).
 RL [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87218454; PubMed=3034567;
 RA Luedin C., Hofstetter H., Sarfati M., Levy C.A., Suter U., Alaimo D., Klicheff E., Frost H., Delespesse G.;
 RT "Cloning and expression of the cDNA coding for a human lymphocyte IgE receptor.";
 RT EMBO J. 6:109-114(1987).
 RL [4]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE=93038513; PubMed=1417742;
 RA Rose K., Turcatti G., Graber P., Pochon S., Regamey P.-O., Jansen K.U., Magnenat E., Audonney N., Bonnetoy J.-Y.;
 RT "Partial characterization of natural and recombinant human soluble CD23.";
 RT Biochem. J. 286:819-824(1992).
 RL [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=89028672; PubMed=2972386;
 RA Yokota A., Kikuchi H., Tanaka T., Sato R., Barsumian E.L., Suenura M., Kishimoto T.;
 RT "Two species of human Fc epsilon receptor II (Fc epsilon RII/CD23): tissue-specific and IL-4-specific regulation of gene expression.";
 RT Cell 55:611-618(1988).
 RL [6]
 RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.
 RX MEDLINE=94191542; PubMed=8142907;
 RA Padlan E.A., Helm B.A.;
 RT "Modeling of the lectin-homology domains of the human and murine low-affinity Fc epsilon receptor (Fc epsilon RII/CD23).";
 RT Receptor 3:325-341(1993).
 RL [7]
 RP 3D-STRUCTURE MODELING OF 173-285.
 RX MEDLINE=96276216; PubMed=8745401;
 RA Bajorath J., Aruffo A.;
 RT "Structure-based modeling of the ligand binding domain of the human cell surface receptor CD23 and comparison of two independently derived molecular models.";
 RT Protein Sci. 5:240-247(1996).
 RL -I- FUNCTION: THIS RECEPTOR HAS ESSENTIAL ROLES IN THE REGULATION OF IGE PRODUCTION AND IN THE DIFFERENTIATION OF B-CELLS (IT IS A B-CELL-SPECIFIC ANTIGEN).

CC -I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS
 CC A SOLUBLE EXCRETED FORM.
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- PTM: N- AND O-GLYCOSYLATED.
 CC MISCELLANEOUS: THERE ARE TWO KINDS OF FC RECEPTORS FOR IGE, WHICH
 CC DIFFER IN BOTH STRUCTURE AND FUNCTION. HIGH AFFINITY RECEPTORS ON
 CC BASOPHILS AND MAST CELLS AND LOW AFFINITY RECEPTORS ON LYMPHOCYTES
 CC AND MONOCYTES.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- DATABASE: NAME=PROW; NOTE=CD guide CD23 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd23.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M15059; AAA52434.1; -;
 DR EMBL: M14766; AAA52435.1; -;
 DR EMBL: X04772; CAA28465.1; -;
 DR EMBL: M23562; AAA52433.1; -;
 DR PIR: A26067; INHDER.
 DR PIR: A26164; A26164.
 DR PIR: A26589; A26589.
 DR PIR: A31924; A31924.
 DR PIR: J10132; J10132.
 DR PDB: 1HLI; 31-JAN-94.
 DR PDB: 1KTE; 03-APR-96.
 DR MIM: 151445; -;
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 KW IGE-binding protein; Transmembrane; Glycoprotein; Receptor; B-cell;
 KW Repeat; Lectin; Signal-anchor; Alternative splicing; 3D-structure.
 FT CHAIN 1 321 MEMBRANE BOUND FORM.
 FT CHAIN 150 321 SOLUBLE FORM.
 FT DOMAIN 1 21 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 22 47 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT DOMAIN 48 321 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 162 284 C-TYPE LECTIN (LONG FORM).
 FT SITE 149 150 CLEAVAGE.
 FT REPEAT 69 89
 FT REPEAT 90 110
 FT REPEAT 111 131
 FT REPEAT 160 288
 FT DISULFID 163 174 BY SIMILARITY.
 FT DISULFID 191 282 BY SIMILARITY.
 FT DISULFID 259 273
 FT CARBOHYD 63 63
 FT VARSPLIC 1 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 269 269 MEGOYS -> MNP5Q (IN ISOFORM B).
 FT CONFLICT 269 269 N -> T (IN REF. 3).
 SQ SEQUENCE 321 AA; 36468 MW; F66708C0B6515B87 CRC64;
 Query Match 23.2%; Score 154.5; DB 1; Length 321;
 Best Local Similarity 32.2%; Pred. No. 1.1e-08;
 Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;
 Oy 3 YHSEKPNMQRARFCNDNTDVAIONKAEIYLEKTLFPSSSYWIGIKIG--GIW 60
 Db 175 YFFGKGTQWVHARACDMEGOLVSHSPEDFLTKHSHSGS--WIGLRINDLKGEF 232
 Oy 61 TWGVTGKSLTEAEENWGDEPNKKNKEDCEIYIKNNKDAKGKNDAC--HKILKAALC 117
 Db 233 IWDGSH---VDYSNWAAGEFTSRSGEDCVMM-----RSGSRWNDAFCDRKLGAWVC 282

RESULT	23
MMGL_MOUSE	
ID	MMGL_MOUSE STANDARD; PRT: 304 AA.
AC	P49300;
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	01-NOV-1997 (Rel. 35, Last annotation update)
DE	Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (macrophage galactose/N-acetylgalactosamine-specific lectin) (MGML).
GN	MGML.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10990;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C3H/HEN;
RX	MEDLINE=92268032; PubMed=1587794;
RA	Sato M., Kawakami K., Osawa T., Toyoshima S.:
RT	"Molecular cloning and expression of cDNA encoding a galactose/N-
RT	acetylglucosamine-specific lectin on mouse tumoricidal
RT	macrophages.";
RL	J. Biochem. 111:331-336(1992).
RL	[2]
RP	SEQUENCE OF 102-120 AND 137-151.
RC	STRAIN=C3H/HEN;
RX	MEDLINE=89197865; PubMed=3241002;
RA	Oda S., Sato M., Toyoshima S., Osawa T.:
RT	"Purification and characterization of a lectin-like molecule specific
RT	for galactose/N-acetyl-galactosamine from tumoricidal macrophages.";
RL	J. Biochem. 104:600-605(1988).
CC	-1- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE
CC	UNITS. MAY PARTICIPATE IN THE INTERACTION BETWEEN TUMORICIDAL
CC	MACROPHAGES AND TUMOR CELLS.
CC	-1- SUBUNIT: HOMO-OLIGOMER.
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein.
CC	-1- TISSUE SPECIFICITY: IS EXPRESSED ON THE SURFACE OF ACTIVATED
CC	MACROPHAGES.
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC	or send an email to license@sdb-sib.ch).
CC	-----
DR	EMBL; S36676; AAB22171.1; .
DR	HSSP; P06734; IKJE.
DR	MGI; MG1:96975; MGI.
DR	InterPro: IPR001304; Lectin_C.
DR	Pfam; PF00059; Lectin_C; 1.
DR	SMART; SM00034; CLECT; 1.
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR	PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
KW	Lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.
FT	DOMAIN 1 35
FT	TRANSEM 36 56
FT	(POTENTIAL).
FT	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT	(POTENTIAL).
FT	EXTRACELLULAR (POTENTIAL).
FT	C-TYPE LECTIN (LONG FORM).
FT	BY SIMILARITY.
FT	DISULFID 201 296
FT	BY SIMILARITY.
FT	DISULFID 274 288
FT	BY SIMILARITY.
FT	CARBOHYD 74 74
FT	N-LINKED (GLCNAG. . .) (POTENTIAL).
FT	CARBOHYD 166 166
FT	N-LINKED (GLCNAG. . .) (POTENTIAL).
SO	SEQUENCE 304 AA; 34596 MM; 3E79CD12C34F5BC CRC64;
Query Match	23.0%; Score 153.5; DB 1; Length 304;
Best Local Similarity	29.5%; Pred. NO. 1.3e+08;

Matches	38:	Conservative	18:	Mismatches	52:	Indels	21:	Gaps	6:
OY	1	WTYH-----YSEKMNQARRFGRDNTDLVALQNKAEIYLEKTLPEFSYTWIGR							54
Db	177	WTEHGSCYWESESEKSPMEADKRYLRENSHLVYVNSLEEQNLNRL--AAVWSIGLT							234
OY	55	KIGGIWTV-GTNSKSLTEAEWNGGEPNN-----KKNKEDVEEYIKRNKDGKWNDA							108
Db	235	DONGWNRWVDGID--FEKGFKWMLALQPDNWGHGLGGEGDCAH1-----TTGGFWNDY							287
OY	109	CHKLKALC 117							
Db	288	CORFERWIC 296							
RESULT 24									
KUCR_MOUSE									
ID	KUCR_MOUSE	STANDARD:	PRT:	548 AA.					
AC	P70194:								
DT	01-NOV-1997 (Rel. 35, Created)								
DT	01-NOV-1997 (Rel. 35, Last sequence update)								
DT	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	Kupfer cell receptor.								
GN	KCLR.								
OC	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxId=10090;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=BALB/C; TISSUE=Liver;								
RA	Takezawa R., Wagatsuma H., Nomoto C., Matanabe Y., Aikai T.;								
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.								
CC	-1 FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.								
CC	COULD BE INVOLVED IN ENDOCYTOSIS.								
CC	-1 SUBCELLULAR LOCATION: Type II membrane protein.								
CC	-1 TISSUE SPECIFICITY: KUPFER CELLS.								
CC	-1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@isb-sib.ch).								
CC	-----								
DR	EMBL; D88577; BAA13647.1; -								
DR	HSSP; P20693. 1HLJ.								
DR	MED; MGI:1859834; KCLR.								
DR	InterPro; IPR001304; lectin_c.								
DR	Pfam; PF00059; lectin_c; 1.								
DR	SMART; SM00034; CLECT; 1.								
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.								
DR	PROSITE; PS50041; C-TYPE_LECTIN_2; 1.								
KW	Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;								
KW	Endocytosis.								
FT	DOMAIN	1	42						
FT	TRANSMEM	43	69						
FT									
FT	DOMAIN	70	548						
FT	DOMAIN	438	538						
FT	DISULFID	440	536						
FT	DISULFID	516	528						
FT	CARBOHYD	86	86						
FT	CARBOHYD	92	92						
FT	CARBOHYD	115	115						
FT	CARBOHYD	132	132						
FT	CARBOHYD	209	209						
FT	CARBOHYD	255	255						
FT	SEQUENCE	548 AA:	61268 MW;	666435E820E73BD9	CNC64;				

Query Match 23.0%; Score 153; DB 1; Length 548;
 Best Local Similarity 30.7%; Pred. No. 2.9e-08;
 Matches 35; Conservative 21; Mismatches 42; Indels 16; Gaps 5;

QY 3 YHSEKPMNORARFRCDNYTDVLAIONKAELEYL-EKTLPSRSYWGIRKIGWT 61
 DB 424 YFESDKPFWREAEKFCYSGAHLASVTSQEOAFVQTT--SSGDHIGLIDQCTEGIW 481
 QY 61 TWV-STNKSLEAEENMGDEPNKN---KEDCVELYIKRNKDGAKWMDAC 109
 DB 482 RWDGTPENNMQSKGFWKGNPDNRHNGREDCVHV-----RQMDNDAC 528

RESULT 25
 LECT_HUMAN STANDARD; PRT; 311 AA.
 ID LECT_HUMAN P07307;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 GN Asialoglycoprotein receptor 2 (Hepatic lectin H2) (ASGP-R) (ASGPR).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=86016723; PubMed=3863106;
 RA Splies M., Lodish H.F.;
 RT "Sequence of a second human asialoglycoprotein receptor: conservation of two receptor genes during evolution."
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6465-6469(1985).
 CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE UNITS. AFTER LIGAND BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE SURFACE.
 CC -1- SUBUNIT: THE FUNCTIONING LIGAND-BINDING UNIT OF THIS RECEPTOR IS THOUGHT TO BE AT LEAST A DIMER.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL CELLS.
 CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M1025; AAB59519.1; -
 CC PIR: A25179; LNH02A.
 CC HSSP: P06734; 1KTE.
 CC MIM: 108361; -
 CC InterPro: JPRO01304; lectin_c.
 CC Pfam: PF00059; lectin_c; 1.
 CC SMART: SM00034; CLECT_1.
 CC PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
 CC PROSITE: PS00641; C_TYPE_LLECTIN_2; 1.
 CC Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
 CC Calcium; Signal-anchor; Phosphorylation.
 CC FT DOMAIN 1 58 CITOPLASMIC (POTENTIAL).
 CC TRANSMEM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC FT (POTENTIAL).
 CC FT DOMAIN 80 311 EXTRACELLULAR (POTENTIAL).
 CC DOMAIN 176 302 C-TYPE LECTIN (LONG FORM).

FT SIZE 5 8 ENDOCYTOSIS SIGNAL (POTENTIAL).
 FT DISULFID 177 188 BY SIMILARITY.
 FT DISULFID 205 300 BY SIMILARITY.
 FT DISULFID 278 292 BY SIMILARITY.
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .).
 FT MOD_RRS 12 12 PHOSPHORYLATION.
 SQ SEQUENCE 311 AA; 35191 MW; 82C78FBC8FBA316 CRC64;

Query Match 22.6%; Score 150.5; DB 1; Length 311;
 Best Local Similarity 30.3%; Pred. No. 2.8e-08;
 Matches 37; Conservative 21; Mismatches 47; Indels 17; Gaps 6;

QY 3 YHSEKPMNORARFRCDNYTDVLAIONKAELEYL-EKTLPSRSYWGIRKIGWT 61
 DB 189 YFESHSKAWAEAEKCYQLENAHLVYINSWEDKFLVQHTNPENT--WIGLIDSGSK 245
 QY 62 WV-STNKSLEAEENMGDEPN---KKNKEDCVELYIKRNKDGAKWMDACIKLKA 115
 DB 246 WVDGTD--YRHNKYNMAVYTPDNMHHGHELDGSEDCVEV-----QPDGRNNDDFCIQVYRW 298
 QY 116 LC 117
 DB 299 VC 300

RESULT 26
 LECB_CHICK STANDARD; PRT; 207 AA.
 ID LECB_CHICK P02707;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hepatic lectin.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91268022; PubMed=2050668;
 RA Bezuska K., Crichlow G.V., Rose J.M., Taylor M.E., Drickamer K.;
 RT "Evolutionary conservation of intron position in a subfamily of genes encoding carbohydrate-recognition domains."
 RL J. Biol. Chem. 266:11604-11609(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88186849; PubMed=3281941;
 RA Mellow T.E., Halberg D., Drickamer K.;
 RT "Endocytosis of N-acetylglucosamine-containing glycoproteins by rat fibroblasts expressing a single species of chicken liver glycoprotein receptor."
 RL J. Biol. Chem. 263:5468-5473(1988).
 RN [3]
 RP SEQUENCE.
 RX MEDLINE=81215504; PubMed=7240175;
 RA Drickamer K.;
 RT "Complete amino acid sequence of a membrane receptor for rat glycoproteins. Sequence of the chicken hepatic lectin."
 RL J. Biol. Chem. 256:5827-5839(1981).
 CC -1- FUNCTION: HEPATIC LECTIN IS A MEMBRANE RECEPTOR PROTEIN THAT RECOGNIZES AND BINDS EXPOSED N-ACETYLGALACTOSAMINE MOIETIES OF PLASMA GLYCOPROTEINS, THUS MEDIATING THEIR CLEARANCE (FROM THE CIRCULATION) AND ENDOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -1- PTM: SOME OR ALL OF THE CYSTEINES ARE INVOLVED IN DISULFIDE BONDS.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63230; AAA48814.1; -.
DR EMBL; M63225; AAA48814.1; JOINED.
DR EMBL; M63226; AAA48814.1; JOINED.
DR EMBL; M63227; AAA48814.1; JOINED.
DR EMBL; M63228; AAA48814.1; JOINED.
DR EMBL; M63229; AAA48814.1; JOINED.
DR EMBL; M63229; AAA48814.1; JOINED.
DR PIR; A03167; LNCNL.
DR HSP; P20693; LNCNL.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
KW Lectin; glycoprotein; Receptor; Endocytosis; Transmembrane;
KW Signal-anchor; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 24 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 1 23 (POTENTIAL).
FT DOMAIN 49 207 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 77 203 C-TYPE LECTIN (LONG FORM).
FT DISULFID 78 92 BY SIMILARITY.
FT DISULFID 109 201 BY SIMILARITY.
FT DISULFID 179 193 BY SIMILARITY.
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 207 AA; 24216 MW; 1F6B36FDB32899DB CRC64;

Query Match 22.4%; Score 149; DB 1; Length 207;
Best Local Similarity 30.4%; Pred. No. 2.5e-08;
Matches 34; Conservative 21; Mismatches 41; Indels 16; Gaps 5;

Oy 3 YHSEKPMNQRARRECDNTDVAIONKAEIYEKTLPPSRSYWIGI--RKIGSIW 60
Db 93 YFFSLSRSMWKAKECEMHSHLIIIDSYAKONFMRTRNER--FWIGLDENGEGEW 150
Oy 61 TWV---GTNKSILTEAEWNMGDEPNPKKNDCEVEIYIKRNKDGAKWDDMC 109
Db 151 QWVDGSTRSSFT----FKRGEPNNKGFNEDCAHW-----TSGQWMDVYC 193

RESULT 27
LECT_MOUSE STANDARD: PRT; 301 AA.
AC P24721;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor 2 (Hepatic lectin 2) (MHL-2) (ASGP-R)
DE (ASGP-R).
GN ASGR2 OR ASGR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=91027942; PubMed=2223888;
RA Sanford J.P., Doyle D.;
RT "Mouse asialoglycoprotein receptor cDNA sequence: conservation of
RT receptor genes during mammalian evolution.";
RL Blochum. Biophys. Acta 1087:259-261(1990).
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
```

```
CC TERMINAL GALACTOSE AND N-ACETYLGLACOSAMINE UNITS. AFTER LIGAND
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC SURFACE.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CELLS.
CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53042; CAA37211.1; -.
DR PIR; S13165; S13165.
DR HSP; P06734; IKGE.
DR MGD; MG1:88082; Asgr2.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
KW Lectin; glycoprotein; Receptor; Endocytosis; Transmembrane;
KW Calcium; signal-anchor; Phosphorylation.
FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 80 301 (POTENTIAL).
FT DOMAIN 169 295 EXTRACELLULAR (POTENTIAL).
FT DISULFID 170 181 C-TYPE LECTIN (LONG FORM).
FT DISULFID 198 293 BY SIMILARITY.
FT DISULFID 271 285 BY SIMILARITY.
FT CARBOHYD 97 97 BY SIMILARITY.
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 301 AA; 34907 MW; 3A29F1AFBA68F298 CRC64;

Query Match 21.5%; Score 143.5; DB 1; Length 301;
Best Local Similarity 28.1%; Pred. No. 1.4e-07;
Matches 34; Conservative 22; Mismatches 50; Indels 15; Gaps 5;

Oy 3 YHSEKPMNQRARRECDNTDVAIONKAEIYEKTLPPSRSYWIGIRKIGSIW 62
Db 182 YWFSRDGLTWAEADQYCOLENAHLVINSREDDVVKHR--SOFHIMIGLDRDGSWKW 239
Oy 63 V-GTNKSILTEAEWNMGDEPNP-----KKNKEDCEVEIYIKRNKDGAKWDDACHRLKAAL 116
Db 240 VDGTG--YRSNVRNMAFLQPDWNGCHEGGGSDCAEL-----SGHNNDRNCGQVNRNV 292
Oy 117 C 117
Db 293 C 293

RESULT 28
LECT_RAT STANDARD: PRT; 301 AA.
AC P08290;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor R2/3 (Hepatic lectin 2/3) (MHL-2) (ASGP-R)
DE (ASGP-R).
GN ASGR2 OR ASGR-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87257885; PubMed=3600647;
 RA McPhaul M., Berg P.;
 RT "Identification and characterization of cDNA clones encoding two
 homologous proteins that are part of the asialoglycoprotein
 receptor.";
 RL Mol. Cell. Biol. 7:1841-1847(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87250656; PubMed=3597443;
 RA Halberg D.F., Wager R.E., Farrell D.C., Hildreth J., Quesenberry M.S.,
 Loeb J.A., Holland E.C., Drickamer K.;
 RT "Major and minor forms of the rat liver asialoglycoprotein receptor
 are independent galactose-binding proteins. Primary structure and
 glycosylation heterogeneity of minor receptor forms.";
 RL J. Biol. Chem. 262:9828-9838(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
 RX MEDLINE=89170119; PubMed=3234178;
 RA Sanford J.P., Elliott R.W., Doyle D.;
 RT "Asialoglycoprotein receptor genes are linked on chromosome 11 in the
 mouse.";
 RL DNA 7:721-728(1988).
 RN [4]
 RP SEQUENCE OF 201-301.
 RX MEDLINE=84111554; PubMed=6319386;
 RA Drickamer K., Mamou J.F., Bins G., Leung J.O.;
 RT "Primary structure of the rat liver asialoglycoprotein receptor.
 Structural evidence for multiple polypeptide species.";
 RL J. Biol. Chem. 259:770-778(1984).
 CC - FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
 WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
 CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
 TERMINAL GALACTOSE AND N-ACETYLGLACTOSAMINE UNITS. AFTER LIGAND
 BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
 TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
 DISSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
 SURFACE.
 CC - SUBCELLULAR LOCATION: TYPE II membrane protein.
 CC - TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
 CELLS.
 CC - MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
 CC - IDENTIFIED: RHU-1 AND RHU-2/3, HAVING A RELATIVE ABUNDANCE OF 4:1.
 CC RHU-2 AND RHU-3 ONLY DIFFERS IN THEIR CARBOHYDRATE STRUCTURES.
 CC - SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 or send an email to license@sdb.ch).
 CC -----
 DR EMBL: M16347; AAA42038.1; -;
 DR EMBL: J02762; AAA41522.1; -;
 DR EMBL: X07636; CAA30476.1; -;
 DR PIR: A25417; INRT2.
 DR PIR: A26888; A26888.
 DR PIR: A28462; A28462.
 DR PIR: A31601; A31601.
 DR HSSP: P06734; 1KJE.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF000059; lectin_c.1.
 DR SMART: SM00034; CLECT.1
 DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
 DR PROSITE: PS00041; C_TYPE_LLECTIN_2; 1.
 KW lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;

KW Calcium; Signal-anchor; Phosphorylation.
 FT DOMAIN 1 58 CITOPLASMIC (POTENTIAL).
 FT TRANSMEM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT DOMAIN 80 301 (POTENTIAL).
 FT DOMAIN 169 295 EXTRACELLULAR (POTENTIAL).
 FT DISULFID 170 181 C-TYPE LECTIN (LONG FORM).
 FT DISULFID 198 293 BY SIMILARITY.
 FT DISULFID 271 285 BY SIMILARITY.
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .).
 FT CONFLICT 153 153 R -> A (IN REF. 1).
 FT CONFLICT 202 202 I -> N (IN REF. 2 AND 3).
 FT CONFLICT 260 260 C -> W (IN REF. 2 AND 3).
 SQ SEQUENCE 301 AA: 34943 MW: 3C2315B642D71279 CRC64;
 Query Match 21.2%; Score 141.5; DB 1; Length 301;
 Best Local Similarity 28.1%; Pred No. 2,2e-07;
 Matches 34; Conservative 19; Mismatches 53; Indels 15; Gaps 5;
 OY 3 YHSEKPMNQRRRCRDNYTDVAIQNKAEILEYDEKTPFSRSYYWIGIRKIGIWTW 62
 DB 182 YWFSRGLTWAEADQYQOMEIAHLIVINSREGEFVVKHR--GAFHIVIGLIDKDGSMKW 239
 OY 63 V-GTKNSLTFEAEKNGWDGEPNN-----KKNKECVETIYIKRKNDAKRWDACHKLKAAL 116
 DB 240 VDSGT--EYNSNFKMMAFPTDPDNCQHEGSGEDCAEIL-----SDGLWMDNCCQOVNRWA 292
 OY 117 C 117
 DB 293 C 293
 RESULT 29
 ID FCE2_MOUSE STANDARD; PRT: 331 AA.
 AC P20693; 061556; 061557;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Low affinity immunoglobulin epsilon FC receptor (lymphocyte IGE
 receptor) (Fc-epsilon-R1) (CD23).
 GN FCE2 OR FCE2A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90017519; PubMed=2529542;
 RA Bettler B., Hofstetter H., Rao M., Yokoyama W.M., Klichherr F.,
 Conrad D.H.;
 RT "Molecular structure and expression of the murine lymphocyte low-
 affinity receptor for IGE (Fc epsilon R1).";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7566-7570(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90171598; PubMed=2137845;
 RA Collinick S.O., Tronstine M.L., Yamashita L.C., Kehy M.R.,
 Moore K.W.;
 RT "Isolation, characterization, and expression of cDNA clones encoding
 the mouse Fc receptor for IGE (Fc epsilon R1).";
 RL J. Immunol. 144:1974-1982(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (FORMS B AND C).
 RC STRAIN=DBA/2;
 RX MEDLINE=94372613; PubMed=8086828;
 RA Kondo H., Ichikawa Y., Nakamura K., Tsuchiya S.;
 RT "Cloning of cDNAs for new subtypes of murine low-affinity Fc receptor
 for IGE (Fc epsilon R1/CD23).";
 RL Int. Arch. Allergy Immunol. 105:38-48(1994).


```

Db 187 YWTSQSGKPEADKTCQLSNNLTVVNSLAEONFLQTHM--GSVTVWIGLTPDONGFWRW 244
QY 63 V-CTN--KSLTEBAENMGDEPNPN-----KKNKEDVEIYIKRNKDGAKMNDACHRLKA 114
Db 245 VDGTDVEKGF-----HWAPRQPDNMYGHGLGGEDCAHF-----TSDGRNNDVYCCQPR 295
QY 115 ALC 117
Db 296 WVC 298

RESULT 31
LECG_TRIST STANDARD; PRT: 158 AA.
ID LECG_TRIST
AC 09YGP1;
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Galactose-binding lectin precursor (TSL).
OS Trimeresurus stejnegeri (Chinese green tree viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=39682;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 24-53, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=99348038; PubMed=10417338;
RA "Zeng R., Xu Q., Shao X.-X., Wang K.-Y., Xia Q.-C.;
RT "Characterization and analysis of a novel glycoprotein from snake
RT venom using liquid chromatography-electrospray mass spectrometry and
RT Edman degradation."
RL Eur. J. Biochem. 266:352-358(1999).
CC - FUNCTION: GALACTOSE-BINDING PROTEIN WHICH RECOGNIZES SPECIFIC
CC CARBOHYDRATE STRUCTURES AND AGGLUTINATE A VARIETY OF ANIMAL CELLS
CC BY BINDING TO CELL-SURFACE GLYCOPROTEINS AND GLYCOLIPIDS. MAY BE A
CC CALCIUM-DEPENDENT LECTIN.
CC - SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC - MASS SPECTROMETRY: MW=17924.2; MW_ERR=2.4; METHOD=Electrospray;
CC RANGE=24-158.
CC - SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF119097; AAD17252.1;
DR HSSP: P05451; ILIT.
DR GlycoSuiteDB: Q9YGP1;
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECN; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR Lectin; Signal; Calcium; Glycoprotein.
KM SIGNAL 1 23
FT CHAIN 24 158 GALACTOSE-BINDING LECTIN.
FT DOMAIN 24 158 C-TYPE LECTIN (LONG FORM).
FT MOD_RES 33 33 OXIDIZED.
FT DISULFID 26 37 BY SIMILARITY.

```

```

FT DISULFID 54 154 BY SIMILARITY.
FT DISULFID 61 156 BY SIMILARITY.
FT DISULFID 129 146 BY SIMILARITY.
FT DISULFID 109 109 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 28 28 N-LINKED (GLCNAC... ) (HIGH MANNOSE).
SQ SEQUENCE 158 AA; 18635 MW; 2077BC62B7A08BF9 CRC64;

Query Match 20.0%; Score 133; DB 1; Length 158;
Best Local Similarity 30.0%; Pred. No. 8, 1e-07;
Matches 39; Conservative 17; Mismatches 46; Indels 28; Gaps 9;

QY 3 YHSEKPMNQARRCRDNY---TDLVALONKAE---LEYLEKTLPEFSRSY----- 49
Db 38 YKLFDEPKTWEDAEFCR--KYKPGCHLASPHRLAESIDIAEYIS-----DYHKROAEV 89
QY 50 WIGI--RKIGITWVGTNKSLEEADNMGDEPNKKKEDCEVEIYIKRNKDGAKMND 107
Db 90 WIGLLDRKKDFSEW--TDRSCDYL--NMDKNDPDHYKKEFCVELVSLGYH--RMNDQ 144
QY 108 ACHKLKALC 117
Db 145 VCESKNSFLC 154

RESULT 32
KUCR_RAT
ID KUCR_RAT STANDARD; PRT: 550 AA.
AC P10716;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Kupffer cell receptor.
GN KCLR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 83-104.
RX MEDLINE=88227939; PubMed=2836387;
RA Hoyle G.W., Hill R.L.;
RT "Molecular cloning and sequencing of a cDNA for a carbohydrate
RT binding receptor unique to rat Kupffer cells."
RL J. Biol. Chem. 263:7487-7492(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91107689; PubMed=1846367;
RA Hoyle G.W., Hill R.L.;
RT "Structure of the gene for a carbohydrate-binding receptor unique to
RT rat Kupffer cells."
RL J. Biol. Chem. 266:1850-1857(1991).
CC - FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCCOSE.
CC COULD BE INVOLVED IN ENDOCYTOSIS.
CC - SUBCELLULAR LOCATION: Type II membrane protein.
CC - TISSUE SPECIFICITY: KUPFFER CELLS.
CC - SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03734; AAA41472.1;
DR EMBL: M55532; AAA40892.1;
DR PIR: A28166; A28166.
DR PIR: A38674; A38674.
DR HSSP: P20693; IHLL.
DR InterPro: IPR000017; Syntaxin.

```


DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00059; lectin_c.1.
DR SMART; SM00034; CLECT.1.
DR SMART; SM00503; Synn.1.
DR PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C-TYPE_LLECTIN_2; 1.
KW Receptor; Transmembrane; Glycoprotein; Lectin; signal-anchor;
Endocytosis.
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 70 550 (POTENTIAL).
FT DOMAIN 438 538 EXTRACELLULAR (POTENTIAL).
FT DISULFID 440 536 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 516 528 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 550 AA; 61104 MW; 9358A6CF4C306270 CRC64;

Query Match 19.8%; Score 132; DB 1; Length 550;
Best Local Similarity 28.9%; Pred. No. 4.3e-06;
Matches 33; Conservative 21; Mismatches 44; Indels 16; Gaps 6;

QY 3 YHSEKPMNMQARRCRDNTDVAIONKAEIYELEKTLPPSRSYWIGIRKIG--GIW 60
DB 424 YFSSDKSMHEENFCVSGAHNLASVTSQEQAFVQIT--NAVDHWIGLTDQTEGNW 481
DB 482 RWDGTPPDYVQSRREWRKGGPDNMWRHNGEREDCV--HLDG---MMNDMAC 528

RESULT 33
LECH_MOUSE STANDARD; PRT; 283 AA.
AC P34927; Q64363;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor 1 (Hepatic lectin 1) (MHL-1) (ASGP-R)
DE (ASGP-R).
GN ASGRI OR ASGR-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93176818; PubMed=8439566;
RA Takezawa R., Shinzawa K., Watanabe Y., Akaike T.;
RT "Determination of mouse major asialoglycoprotein receptor cDNA
sequence.";
RL Biochim. Biophys. Acta 1172:220-222(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=95047431; PubMed=7958950;
RA Monroe R.S., Huber B.E.;
RT "The major form of the murine asialoglycoprotein receptor: cDNA
sequence and expression in liver, testis and epididymis.";
RL Gene 148:237-244(1994).
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CARBOHYDRATE MOTIFIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE UNITS. AFTER LIGAND
BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE

CC SURFACE.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CELLS.
CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D13517; BAA02734.1; -;
DR EMBL; 009362; AAB60441.1; -;
DR EMBL; 008372; AAB60440.1; -;
DR PIR; S29855; S29855.
DR HSSP; P20693; 1HLJ.
DR MGD; MGI:88081; Asgr1.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00059; lectin_c.1.
DR SMART; SM00034; CLECT.1.
DR PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C-TYPE_LLECTIN_2; 1.
KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
KW Calcium; signal-anchor; Phosphorylation.
FT INT_MET 0 0
FT DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 60 283 (POTENTIAL).
FT DOMAIN 151 277 EXTRACELLULAR (POTENTIAL).
FT SITE 4 7 C-TYPE LECTIN (LONG FORM).
FT DISULFID 152 163 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 180 275 BY SIMILARITY.
FT DISULFID 253 267 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 150 150 I -> T (IN REF. 1).
SO SEQUENCE 283 AA; 32472 MW; 982A5D305AAED08F CRC64;

Query Match 19.4%; Score 129.5; DB 1; Length 283;
Best Local Similarity 27.0%; Pred. No. 3.6e-06;
Matches 33; Conservative 20; Mismatches 52; Indels 17; Gaps 6;

QY 3 YHSEKPMNMQARRCRDNTDVAIONKAEIYELEKTL--PFSRYWIGIRKIGIWT 61
DB 164 YFSSSVRPWTEADKYCDLENANLAVVTSRDQNLQRHMGCLNT--WIGLTDQNGPWK 220
QY 62 WV-GTNKSLTEEAENWMDGEPNN-----KKNEDCEVEIYIKRNKAGKMWNDACHKLKLA 115
DB 221 WVDGTD--YETGFCQWNRBQPDNMWYGHGLGSGEDCAHF-----TTDGRNNDVYCARPRYRW 273
QY 116 LC 117
DB 274 VC 275

RESULT 34
PSPD_MOUSE STANDARD; PRT; 374 AA.
ID PSPD_MOUSE
AC P50404;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
GN SFPD OR SFPD4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6/CBA; TISSUE=Lung;
RX MEDLINE=96094460; PubMed=7499852;
RA Mowat M., White R.A., Guo N., Dowler L.L., Tauber A.I., Sastri K.N.;
RT "Mouse surfactant protein-D, cDNA cloning, characterization, and gene
RL localization to chromosome 14."
RJ J. Immunol. 155:5671-5677(1995).
CC -1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
CC MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
CC EXTENT OTHER ALPHA-GLUCOSYL MOETIES. IT COULD PARTICIPATE IN THE
CC EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L40156; AAA92021.1; -.
DR HSSP: P35247; 1B08.
DR MGD: MGI:109515; Sftpd.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF01391; Collagen; 3.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW Glycoprotein; Calcium; Surface film; gaseous exchange; Hydroxylation;
KW Signal; Lectin; Collagen; Repeat; Coiled coil.
FT SIGNAL 1 19
FT CHAIN 20 374
FT FT PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT D.
FT DOMAIN 45 221
FT FT COLLAGEN-LIKE.
FT DOMAIN 222 253
FT FT COILED COIL (POTENTIAL).
FT DOMAIN 278 374
FT FT C-TYPE LECTIN (SHORT FORM).
FT DISULFID 280 372
FT FT BY SIMILARITY.
FT DISULFID 350 364
FT FT BY SIMILARITY.
FT CARBOHYD 89 89
FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 374 AA; 37688 MW; FE034261263FA3E4 CRC64;

```

```

Query Match 18.9%; Score 126; DB 1; Length 374;
Best Local Similarity 28.1%; Pred. No. 1,1e-05;
Matches 32; Conservative 24; Mismatch 46; Indels 12; Gaps 5;

```

```

OY 6 SEKPMNORARRRCRDNYIDLVAIONKAELEYLEKTLPSRSYWGIRKIG--GIWTWY 63
DB 269 SEKP--FEDAGQEMCKOAGOLASPRASATENAIOQLITAHNKAALFSLMDVGEKFTY- 325
OY 64 GTNKSLTEAEWNGDEPNKKKNKEDCEVETIYIKRNKADGRWMDACHKLKAAL 117
DB 326 PTGEPLV--YSNWAPEGPNNNGAENCVEIFTN-----GQWMDKAGGEGRIYIC 372

```

```

RESULT 35
PSPD_RAT
ID PSPD_RAT STANDARD: PRT; 374 AA.
AC P35248;
DT 01-FEB-1994 (Rel. 28, Created)

```

```

DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D)
DE (CP4).
GN SFTPD OR SFTP4.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-33.
RC TISSUE=Lung;
RX MEDLINE=92112913; PubMed=1370483;
RA Shimizu H., Fisher J.H., Papst P., Benson B., Lau K., Mason R.J.,
RA Voelker D.R.;
RT "Primary structure of rat pulmonary surfactant protein D. cDNA and
RT deduced amino acid sequence."
RJ J. Biol. Chem. 267:1853-1857(1992).
RN [2]
RP SEQUENCE OF 73-95 AND 153-180.
RC TISSUE=Lung;
RX MEDLINE=90001186; PubMed=2675969;
RA Persson A., Chang D., Rust K., Moxley M., Longmore W., Crouch E.;
RT "Purification and biochemical characterization of CP4 (SP-D), a
RT collagenous surfactant-associated protein."
RJ Biochemistry 28:6361-6367(1989).
CC -1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
CC MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
CC EXTENT OTHER ALPHA-GLUCOSYL MOETIES. IT COULD PARTICIPATE IN THE
CC EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M81231; AAA42170.1; -.
DR PIR: A42046; A42046.
DR HSSP: P35247; 1B08.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF01391; Collagen; 3.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW Signal; Lectin; Collagen; Repeat; Coiled coil.
FT SIGNAL 1 19
FT CHAIN 20 374
FT FT PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT D.
FT DOMAIN 45 221
FT FT COLLAGEN-LIKE.
FT DOMAIN 222 253
FT FT COILED COIL (POTENTIAL).
FT DOMAIN 278 374
FT FT C-TYPE LECTIN (SHORT FORM).
FT DISULFID 280 372
FT FT BY SIMILARITY.
FT DISULFID 350 364
FT FT BY SIMILARITY.
FT CARBOHYD 89 89
FT FT N-LINKED (GLCNAC. . .).
FT MOD_RES 77 77
FT MOD_RES 86 86
FT MOD_RES 95 95
FT MOD_RES 98 98
FT MOD_RES 170 170
FT MOD_RES 170 170

```

FT MOD_RES 176 176 HYDROXYLATION.
 FT CONFLICT 89 N -> E (IN REF. 2).
 FT CONFLICT 164 K -> C (IN REF. 2).
 SO SEQUENCE 374 AA; 37561 MM; DB2BB5E39DB4A3C CRC64;

Query Match 18.8%; Score 125; DB 1; Length 374;
 Best Local Similarity 28.1%; Pred. No. 1.4e-05;

Matches 32; Conservative 25; Mismatches 45; Indels 12; Gaps 5;

OY 6 SEKPMNQRARFCRDNTDVAIONKAEILEKTLFSSRSYWGIRKIG-GIWTWV 63
 DB 269 SEEP--FEDAKEMCRQAGQGLASPRATENAVALVTAAHKAAPLSMTDVGTEGKFTY- 325

OY 64 GTNKSLTPEAEWNGDGEPPNKKKEDCEVEIYIKRNKDGAKNDACGHLKAALC 117
 DB 326 PTGEALV--YSNMAPGEPPNNGAGNCEVEITFN-----GQNMNDACGGEORLVIC 372

RESULT 36

MABC_MOUSE STANDARD; PRT: 244 AA.

AC PA1317;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Mannose-binding protein C precursor (MBP-C) (mannan-binding protein)
 DE (RA-reactive factor P28a subunit) (RARF/P28A).
 GN MBL2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=95284466; PubMed=7766991;
 RA Sastiy R., Wang J.S., Brown D.C., Ezekowitz R.A., Tauber A.I.,
 RA Sastiy K.N.;
 RT "Characterization of murine mannose-binding protein genes Mb1 and
 RT Mb2 reveals features common to other collectin genes.";
 RL J. Immunol. 147:692-697(1991).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=95284466; PubMed=7766991;
 RA Sastiy R., Wang J.S., Brown D.C., Ezekowitz R.A., Tauber A.I.,
 RA Sastiy K.N.;
 RT "Characterization of murine mannose-binding protein genes Mb1 and
 RT Mb2 reveals features common to other collectin genes.";
 RL Mamm. Genome 6:103-110(1995).
 RN [3]

RP SEQUENCE FROM N.A.
 RA Kuge S., Ihara S., Watanabe E., Watanabe M., Takishima K., Suga T.,
 RA Mamaiya G., Kawakami M.;
 RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: BINDS MANNOSE AND N-ACETYLGLUCOSAMINE IN A CALCIUM-
 CC DEPENDENT MANNER. IS CAPABLE OF HOST DEFENSE AGAINST PATHOGENS,
 CC BY ACTIVATING THE CLASSICAL COMPLEMENT PATHWAY INDEPENDENTLY OF
 CC THE ANTIBODY.
 CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
 CC -1- SUBCELLULAR LOCATION: THIS PROTEIN IS LOCATED INTRACELLULAR, MOST
 CC PREDOMINANTLY IN THE ROUGH ENDOPASMIC RETICULUM AND IN THE GOLGI
 CC APPARATUS, AND LITTLE OR NO MBP IS PRESENT IN PLASMA MEMBRANES AND
 CC LYSOSOMES. THE MBP IS EXCLUSIVELY LOCALIZED IN THE CISTERNA
 CC SPACE OF THE ORGANELLES, PROBABLY AS A LOOSELY BOUND MEMBRANE
 CC PROTEIN.

CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; S42294; AAB19343.1; -;
 DR EMBL; U09016; AAB82010.1; -;
 DR EMBL; U09013; AAB82010.1; JOINED.
 DR EMBL; U09014; AAB82010.1; JOINED.
 DR EMBL; U09015; AAB82010.1; JOINED.
 DR EMBL; D11440; BAA02005.1; -;
 DR HSSP; P08661; IRDO.
 DR MGD; MGI:96924; Mb12.

DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam; PF00059; lectin_c_1.

DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
 DR PROSITE; PS00041; C-TYPE_LLECTIN_2; 1.

KW Lectin; Hydroxylation; Liver; Glycoprotein; Mannose-binding; Membrane;
 KW Calcium; Collagen; Repeat; Signal.

FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 244 MANNOSE-BINDING PROTEIN C.
 FT DOMAIN 38 96 COLLAGEN-LIKE.
 FT DOMAIN 149 242 C-TYPE LECTIN (SHORT FORM).
 FT MOD_RES 43 43 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 58 58 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 69 69 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 78 78 HYDROXYLATION (POTENTIAL).
 FT MOD_RES 81 81 HYDROXYLATION (POTENTIAL).
 FT DISULFD 29 29 INTERCHAIN (BY SIMILARITY).
 FT DISULFD 34 34 INTERCHAIN (BY SIMILARITY).
 FT DISULFD 151 240 INTERCHAIN (BY SIMILARITY).
 FT DISULFD 218 232 BY SIMILARITY.

FT CONFLICT 3 3 I -> L (IN REF. 1).
 FT CONFLICT 15 15 V -> A (IN REF. 1).
 SO SEQUENCE 244 AA; 25957 MM; 49AE84E2290DEBDA CRC64;

Query Match 18.5%; Score 123.5; DB 1; Length 244;
 Best Local Similarity 29.6%; Pred. No. 1.3e-05;
 Matches 34; Conservative 14; Mismatches 52; Indels 15; Gaps 4;

OY 6 SEKPMNQRARFCRDNTDVAIONKAEILEKTLFSSRSYWGIRKIG-GIWTWV 62
 DB 138 SVKXSLDRVAKALCEFGSVATPRNAENSAIQV---ADIAVLGITDVAVESSEFEDL 194
 OY 63 VGTNKSLTPEAEWNGDGEPPNKKKEDCEVEIYIKRNKDGAKNDACGHLKAALC 117
 DB 195 TGNRYRYT-----NMWDGEPNNTGDEDCVIL-----GNGKRWNDVPCSDSFLAIC 240

RESULT 37

LEC3_MEGRO STANDARD; PRT: 162 AA.

AC P07439;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lectin BRA-3 precursor.
 OS Megabalanus rosa (Acorn barnacle).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;
 OC Thoracica; Sessilia; Balanomorpha; Balanoidae; Balanidae;
 OC Megabalanus.

OX NCBI_TaxID=6680;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93292994; PubMed=8514190;
 RA Takamatsu N., Takeda T., Kojima M., Helshi M., Muramoto K.,
 RA Kamuya H., Shiba T.;

RT "Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene
 RT structure and seasonal changes of mRNA and protein levels.";
 RL Gene 128:251-255(1993).
 RN [2]

```

RP SEQUENCE OF 25-162.
RA Muramoto K., Kamuya H.;
RT "The amino-acid sequence of a lectin of the acorn barnacle
RL Megabalanus rosa."
CC Biochim. Biophys. Acta 874:285-295(1986).
CC -1- FUNCTION: SUGAR-BINDING PROTEIN WHICH RECOGNIZES SPECIFIC
CC CARBOHYDRATE STRUCTURES AND AGGLOMERATE A VARIETY OF ANIMAL
CC CELLS BY BINDING TO CELL-SURFACE GLYCOPROTEINS AND GLYCOLIPIDS.
CC THIS IS A CALCIUM-DEPENDENT LECTIN. INVERTEBRATE LECTINS MAY BE
CC INVOLVED IN DEFENSE FUNCTIONS.
CC -1- SUBUNIT: HOMOTETRAMER. DISULFIDE-LINKED.
CC -1- TISSUE SPECIFICITY: COELOMIC FLUID.
CC -1- MISCELLANEOUS: THIS LECTIN BINDS GALACTOSE.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D13302; BAA02556.1; -
DR EMBL: D13298; BAA02556.1; JOINED.
DR EMBL: D13300; BAA02556.1; JOINED.
DR EMBL: D13301; BAA02556.1; JOINED.
DR PIR: A26094; LNR3.
DR HSSP: P23806; LNR3.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_2; 1.
DR KWLectin; Calcium; Signal.
FT SIGNAL 1 24
FT CHAIN 25 162 LECTIN BRA-3.
FT DOMAIN 25 152 C-TYPE LECTIN (LONG FORM).
FT DISULFID 26 39
FT DISULFID 56 150
FT DISULFID 125 142
FT DISULFID 157 157 INTERCHAIN (WITH C-136 IN OTHER CHAIN).
FT DISULFID 160 160 INTERCHAIN (WITH C-133 IN OTHER CHAIN).
FT VARIANT 146 146 K -> R.
SQ SEQUENCE 162 AA; 18328 MW; EBF14E91DD1CB81 CRC64;

```

Query Match 18.3%; Score 122; DB 1; Length 162;
 Best Local Similarity 30.6%; Pred. No. 1.1e-05;
 Matches 37; Conservative 21; Mismatches 47; Indels 16; Gaps 8;

```

DE (Islet cells regeneration factor) (ICRF).
GN REG1 OR REG.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91093273; PubMed=1985964;
RA Rouquier S., Verdier J.M., Iovanna J., Dagorn J.-C., Giorgi D.;
RT "Rat pancreatic stone protein messenger RNA. Abundant expression in
RT mature exocrine cells, regulation by food content, and sequence
RT identity with the endocrine reg transcript."
RL J. Biol. Chem. 266:786-791(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88115343; PubMed=2963000;
RA Terazono K., Yamamoto H., Takasawa S., Shiga K., Yonemura Y.,
RA Tochino Y., Okamoto H.;
RT "A novel gene activated in regenerating islets."
RL J. Biol. Chem. 263:2111-2114(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93326645; PubMed=7916640;
RA Dusetli N.J., Frigerio J.M., Dagorn J.-C., Iovanna J.L.;
RT "Rapid PCR cloning and sequence determination of the rat
RT 11thosathine gene."
RL Biochim. Biophys. Acta 1174:99-102(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR;
RA Miyashita H., Suzuki Y., Watanabe T., Uno M., Morizumi S.,
RA Yonekura H., Okamoto H.;
RT "Structure and characterization of rat Reg I gene."
RL Selkagaku 65:1082-1082(1993).
RN [5]
RP SEQUENCE OF 22-69.
RC TISSUE=Pancreas;
RX MEDLINE=90031455; PubMed=2680252;
RA Adrich Z., de Caro A.M., Guidoni A.A., Woudstra M.E., Rovey M.;
RT "Characterization in rat pancreatic juice of a protein homologous to
RT the human pancreatic stone protein."
RL Comp. Biochem. Physiol. 93B:793-797(1989).
CC -1- FUNCTION: MIGHT ACT AS AN INHIBITOR OF SPONTANEOUS CALCIUM
CC CARBONATE PRECIPITATION.
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN REGENERATING ISLETS, BUT
CC NOT IN NORMAL PANCREATIC ISLETS, INSULINOMAS OR REGENERATING
CC LIVER.
CC -----
CC -1- SIMILARITY: TO OTHER MEMBERS OF THE C-TYPE LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L07512; AAA41533.1; -
DR EMBL: M62830; AAA41974.1; -
DR EMBL: M18962; AAA42028.1; -
DR EMBL: D26164; BAA05149.1; -
DR PIR: A28351; A28351.
DR PIR: PL0147; PL0147.
DR PIR: A39081; A39081.
DR PIR: S34618; S34618.
DR HSSP: P03451; ILIT.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR003990; pancreatitis_assoc.
DR Pfam: PF00059; lectin_c; 1.
DR PRINTS: PR01504; PNCREATTISAP.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.

```



```

FT DISULFID 2010 2026 BY SIMILARITY.
FT DISULFID 2041 2084 BY SIMILARITY.
FT DISULFID 2070 2097 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1842 1842 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 2124 AA; 221117 MW; E30BB6153A34B1 CRC64;

Query Match 18.2%; Score 121; DB 1; Length 2124;
Best Local Similarity 31.7%; Pred. No. 0.00027;
Matches 38; Conservative 17; Mismatches 45; Indels 20; Gaps 8;

Qy 1 WT-----YHSEKPMNQARRRCRDNYDVAIQNAETLEYEKLTPSRSYWIGI- 53
Db 1918 WTKFQGHCHYRFPDPRETWVAERCRQOSHLSIVTPEEQEFVNKN---AQDYQWIGLN 1974

Qy 54 -RKIGIWTWYGTNKSILTEAENMGDEPPNN-KNKEDC-VEIYIKNKNDAGKNDNDACH 110
Db 1975 DRTIEGDFRW-SDGHSL--QEKRRPNQDPNFEATGECVVMWHER---GEMNDVPCN 2027

RESULT 40
PGCA_MOUSE STANDARD; PRT; 2132 AA.
ID PGCA_MOUSE
AC 0612B2; 064021;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core
  protein) (CSPC).
GN AGC1 OR AGC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=Cartilage;
RX MEDLINE=95104847; PubMed=7806222;
RA Walcz E., Deak F., Ethardt P., Coulter S.N., Fueleop C., Horvath P.,
  Doege K.J., Giant T.T.;
RT "Complete coding sequence, deduced primary structure, chromosomal
  localization, and structural analysis of murine aggrecan.";
RL Genomics 22:364-371(1994).
RN (2)
RP SEQUENCE OF 211-326 FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=95004579; PubMed=7920633;
RA Watanabe H., Kimata K., Line S., Strong D., Gao L.-Y., Kozak C.A.,
  Yamada Y.;
RT "Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in
  the aggrecan gene.";
RL Nat. Genet. 7:154-157(1994).
CC -1- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
  MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
  IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
  HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A
  REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
  SIMILARITY).
CC -1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
  TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
  MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
  CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
  THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
  AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
  AND G3.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE

```

```

CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
CC -1- DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF CARTILAGE MATRIX
  DEFICIENCY (CMD), AN AUTOSOMAL RECESSIVE SYNDROME CHARACTERIZED BY
  CLEFT PALATE, SHORT LIMBS, TAIL AND SNOUT. MUTATION IN STRAIN CMD
  CAUSES ABSENCE OF AGGREGAN BY TRUNCATION OF THE PROTEIN (MUTATION
  IN THE G1 DOMAIN).
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 4 LINK DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sdb.ch/announce/
  or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: L07049; AAC37670.1; -.
DR EMBL: S73722; AAB32160.1; -.
DR EMBL: S73721; AAB32160.1; JOINED.
DR HSSP: P98066; ITSG.
DR MGD: MGI:99602; Agc.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR InterPro: IPR000558; Link.
DR InterPro: IPR003324; SGXSG.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00059; lectin_c.1.
DR Pfam: PF02339; lectin_c.2.
DR Pfam: PF00084; sush1.1.
DR Pfam: PF00193; xlink.4.
DR ProDom: PD000918; Link; 4.
DR SMART: SM00032; CCP.1.
DR SMART: SM00034; CLECT.1.
DR SMART: SM00406; IGV.1.
DR SMART: SM00445; LINK.4.
DR PROSITE: PS00290; IG_MHC.1.
DR PROSITE: PS01241; LINK.4.
DR PROSITE: PS00615; C-TYPE LECTIN.1; 1.
DR PROSITE: PS50041; C-TYPE LECTIN.2; 1.
KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sush1;
KW Repeat; Immunoglobulin domain.
FT SIGNAL 1 19
FT CHAIN 20 2132
FT DOMAIN 44 140 AGGREGAN CORE PROTEIN.
FT DOMAIN 170 247 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 268 349 LINK 1.
FT DOMAIN 504 581 LINK 2.
FT DOMAIN 602 683 LINK 3.
FT DOMAIN 1918 2044 LINK 4.
FT DOMAIN 2048 2106 C-TYPE LECTIN.
FT DOMAIN 48 140 SUSHI.
FT DOMAIN 152 247 G1-B.
FT DOMAIN 253 349 G1-B'.
FT DOMAIN 486 580 G2-B.
FT DOMAIN 587 682 G2-B'.
FT DOMAIN 805 803 KS.
FT DOMAIN 805 1231 CS-1.
FT DOMAIN 1232 1917 CS-2.
FT DOMAIN 1917 2132 G3.
FT DISULFID 51 133 BY SIMILARITY.
FT DISULFID 175 246 BY SIMILARITY.
FT DISULFID 199 220 BY SIMILARITY.
FT DISULFID 273 348 BY SIMILARITY.
FT DISULFID 297 318 BY SIMILARITY.
FT DISULFID 509 580 BY SIMILARITY.
FT DISULFID 533 554 BY SIMILARITY.
FT DISULFID 607 682 BY SIMILARITY.

```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:12:02 : Search time 36.49 Seconds
(without alignments)
308.097 Million cell updates/sec

Title: US-09-119-209-2_COPY_39_155

Perfect score: 666

Sequence: 1 WYHYSEKPMNQRARRRCR.....NKDAGKWNDDACHKLAALC 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	323	1 S09702	L-selectin precurs
2	666	100.0	385	1 A34015	L-selectin precurs
3	631	94.7	372	2 J05377	L-selectin precurs
4	598	89.8	376	2 J04892	L-selectin precurs
5	581	87.2	370	2 S2124	L-selectin precurs
6	565	84.8	372	1 A32375	L-selectin precurs
7	557	83.6	372	2 A23936	L-selectin precurs
8	455	68.3	768	2 A42755	P-selectin precurs
9	451	67.7	646	2 J06473	P-selectin precurs
10	445	66.8	485	2 S36772	P-selectin - bovin
11	445	66.8	830	2 A30359	P-selectin precurs
12	442	66.4	768	2 I53821	P-selectin - rat
13	434	65.2	551	2 I46709	endothelial leukoc
14	425	63.8	482	2 J05092	E-selectin - pig
15	415	62.3	610	2 A35046	E-selectin precurs
16	405	60.8	612	2 B42755	E-selectin precurs
17	161.5	24.2	404	2 A46274	HIV gp120-binding
18	157	23.6	1456	1 A36563	Mannose receptor p
19	154.5	23.2	321	1 LNHURR	IgE Fc receptor II
20	153.5	23.0	304	2 JX0209	lectin, galactose/
21	150.5	22.6	311	1 LNH02A	asialoglycoprotein
22	149	22.4	207	1 LNH02A	hepatic lectin - c
23	148	22.2	1479	2 T42710	mannose receptor,
24	144	21.6	1455	1 A48925	mannose receptor p
25	143.5	21.5	301	2 S13155	asialoglycoprotein
26	142.5	21.4	309	1 S34198	IgE Fc receptor II
27	141.5	21.2	331	1 LNM5ER	IgE Fc receptor, I
28	140.5	21.1	301	1 LNMRT2	hepatic lectin 2 -
29	136.5	20.5	742	2 J07595	scavenger receptor

30	133.5	20.0	306	2 A42230	lectin M-ASGP-BP p
31	132	19.8	550	2 A28166	Kupffer cell recep
32	131.5	19.7	155	2 S78774	perlecan - Halioti
33	129.5	19.4	284	2 S29855	asialoglycoprotein
34	125	18.8	374	1 A42046	surfactant protein
35	123.5	18.5	244	1 LNM5MC	mannose-binding le
36	122	18.3	162	1 LNRG3	lectin BRA3-2 prec
37	121	18.2	165	2 A28351	pancreatic stone p
38	121	18.2	2124	2 A28452	proteoglycan core
39	121	18.2	2132	1 A55182	aggreacan precursor
40	120.5	18.1	284	1 LNMRTL	hepatic lectin - r
41	119.5	17.9	301	2 A53570	collactin-43 - bov
42	119	17.9	162	1 LNRG1	lectin BRA3-1 prec
43	119	17.9	321	2 T26152	hypothetical prote
44	118.5	17.8	2415	1 A39086	aggreacan precursor
45	118	17.7	3562	2 A47171	chondroitin sulfat

ALIGNMENTS

```
RESULT 1
S09702
L-selectin precursor, short splice form - human
N:Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukoc
tal lymph node homing receptor Leu-8
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 26-May-2000
C:Accession: S09702
R:Cammerini, D.; James, S. P.; Stamenkovic, I.; Seed, B.
Nature 342, 78-82, 1998
A:Title: Leu-8/TO1 is the human equivalent of the Mel-14 lymph node homing receptor.
A:Reference number: S06798; MUID:90044046
A:Accession: S09702
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-323 <CAM>
A:Cross-references: EMBL:X17519; NID:q34344
A>Note: this translation is not annotated in GenBank entry HSLEU8, release 111.0
C:Comment: For an alternative splice form, see PIR:A34015.
C:Gene: GDB:SEL; GDB:LNHR; LSEL; LAM1; LYAM1; LAM-1
A:Cross-references: GDB:120157; GDB:118834; OMIM:153240
A:Map position: 1q22-1q23
A:Introns: 14/3; 42/1; 171/1; 207/1; 269/1; 303/2
C:Function:
A:Description: binds with low affinity to oligosaccharides like heparan sulfate and si
ment of leukocytes to areas of inflammation, and with CD162 mediates neutrophil-neutr
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology
C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; Inflammat
F:1-51/Domain: signal sequence #status predicted <SIG>
F:42-168/Domain: C-type lectin homology <LCH>
F:52-300/Domain: extracellular #status predicted <EXT>
F:173-204/Domain: EGF homology <EGF>
F:210-267/Domain: complement factor H repeat homology <FHN>
F:272-323/Domain: complement factor H repeat homology #status atypical <FN2>
F:301-318/Domain: transmembrane #status predicted <TRM>
F:319-333/Domain: intracellular #status predicted <INT>
F:73,117,190,245,259/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:322/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 666; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRARRCNDYDVLVAIONKAELEYEKLTPFSRSYWGIRKIGTW 60
DB 52 WYHYSEKPMNQRARRCNDYDVLVAIONKAELEYEKLTPFSRSYWGIRKIGTW 111
QY 61 TWGVTNKSLEBEAENWGDGEPNNKKNKEDCEVETIKRNKDKGKWNDDACHKLAALC 117
|||||
```

Db 112 TWGVTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 168

RESULT 2

A34015

L-selectin precursor, long splice form - human

N:Alternate names: CD62L; leukocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukocyte cell adhesion molecule-1 (LECAM-1, LAM-1)

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: 155333; S06798; J01014; A34015; A33912

R:Ord, D.C.; Enst, T.J.; Zhou, L.J.; Rambaldi, A.; Spertini, O.; Griffin, J.; Tedder, R.

J: Biol. Chem. 263, 7760-7767, 1990

A:Title: Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Leu-8/TO1)

A:Reference number: 155333; MUID:90243637

A:Accession: 155333

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 14-385 <ORD>

A:Cross-references: GB:M32414; NID:g187259; PIDN:AAB60700.1; PID:g386860

R:Camerini, D.; James, S.P.; Stamenkovic, I.; Seed, B.

Nature 342, 78-82, 1989

A:Title: Leu-8/TO1 is the human equivalent of the Mel-14 lymph node homing receptor.

A:Reference number: S06798; MUID:90044046

A:Accession: S06798

A:Molecule type: mRNA

A:Residues: 1-225, 'S', 227-385 <CAM>

A:Cross-references: EMBL:X17519; NID:g34344; PIDN:CA83536.1; PID:g4902829

A:Note: this translation is not annotated in GenBank entry HSLER8, release 111.0

R: Tedder, T.F.; Isaacs, C.M.; Ernst, T.J.; Demetri, G.D.; Adler, D.A.; Distche, C.M.

J: Exp. Med. 170, 123-133, 1989

A:Title: Isolation and chromosomal localization of cDNAs encoding a novel human lymphocyte protein.

A:Reference number: J01014; MUID:89310350

A:Accession: J01014

A:Molecule type: mRNA

A:Residues: 1-230, 'N', 232, 'N', 234-254, 'E', 256-385 <TED>

A:Cross-references: GB:X16150; NID:g34448; PIDN:CA834275.1; PID:g34429

A:Note: the translated sequence in GenBank entry HSLVMI, release 111.0, differs from the J. Cell Biol. 109, 421-427, 1989

J: Cell Biol. 109, 421-427, 1989

A:Title: Characterization of a human homologue of the murine peripheral lymph node homing receptor.

A:Reference number: A34015; MUID:89308881

A:Accession: A34015

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 14-49, 'Y', 51-190, 'H', 192-205, 'L', 207-226, 'F', 228-385 <BOH>

A:Cross-references: GB:X16070; NID:g38092; PIDN:CA834203.1; PID:g38093

R: Siegelman, M.H.; Weissman, I.L.

Proc. Natl. Acad. Sci. U.S.A. 86, 5562-5566, 1989

A:Title: Human homologue of mouse lymph node homing receptor: evolutionary conservation

A:Reference number: A33912; MUID:89315837

A:Accession: A33912

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 14-205, 'L', 207-385 <SIE>

A:Cross-references: GB:M25280; NID:g187182; PIDN:AAC63053.1; PID:g307134

A:Comment: For an alternative splice form, see PIR:S09702.

C:Genetics:

A:Gene: GDB:SEL1; GDB:LNHR; LSEL; LAM1; LYAM1; LAM-1

A:Cross-references: GDB:120157; GDB:116834; OMIM:153240

A:Map position: 1922-1923

A:Intons: 14/3; 42/1; 171/1; 207/1; 269/1; 331/1; 374/1; 380/2

C:Function:

A:Description: binds with low affinity to oligosaccharides like heparan sulfate and sialyl ment of leukocytes to areas of inflammation, and with CD162 mediates neutrophil-neutrophil

C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;

C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammation

F:1-51/Domain: signal sequence #status predicted <SIG>

F:42-168/Domain: C-type lectin homology <LCH>

F:52-385/Product: L-selectin #status predicted <LCH>

F:52-343/Domain: extracellular #status predicted <EXT>

F:173-204/Domain: EGF homology <EGF>

F:210-267/Domain: complement factor H repeat homology <FH1>

F:272-329/Domain: complement factor H repeat homology <FH2>

F:344-368/Domain: transmembrane #status predicted <TM>

F:369-385/Domain: intracellular #status predicted <INT>

F:73,117,190,245,259/Binding site: carbohydrate (asn) (covalent) #status predicted

F:377,380/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 666; DB 1; Length 385;

Best Local Similarity 100.0%; Pred. No. 3.4e-58;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 112 TWGVTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 168

RESULT 3

J03377

L-selectin precursor - hamadryas baboon

C:Species: Papio hamadryas (hamadryas baboon)

C:Date: 02-Jun-1997 #sequence_revision 12-Sep-1997 #text_change 21-Jan-2000

C:Accession: J03377; PC4315

R:Tsunashita, N.; Fu, H.; Berg, E.L.

Gene 181, 219-220, 1996

A:Title: PCR cloning of the cDNA encoding baboon L-selectin.

A:Reference number: J03377; MUID:97128794

A:Accession: J03377

A:Molecule type: mRNA

A:Residues: 1-372 <TSU1>

A:Cross-references: GB:U52074; NID:g1326148; PIDN:AAB40903.1; PID:g1326149

A:Accession: PC4315

A:Molecule type: protein

A:Residues: 37-43;142-148 <TSU2>

C:Comment: This receptor is involved in the initial adhesive interaction between lymph sites of inflammation.

C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology

F:1-38/Domain: signal sequence #status predicted <SIG>

F:29-155/Domain: C-type lectin homology <LCH>

F:39-372/Product: L-selectin #status predicted <KAT>

F:39-157/Domain: calcium-binding #status predicted <CA>

F:160-191/Domain: EGF homology <EGF>

F:197-254/Domain: complement factor H repeat homology <FH1>

F:259-316/Domain: complement factor H repeat homology <FH2>

F:333-335/Domain: transmembrane #status predicted <TM>

F:356-372/Domain: intracellular #status predicted <INT>

Query Match 94.7%; Score 631; DB 2; Length 372;

Best Local Similarity 94.0%; Pred. No. 9.5e-55;

Matches 110; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 112 TWGVTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 168

RESULT 4

J04892

L-selectin precursor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 21-Jan-2000

C:Accession: J04892

R:Qian, J.; Huang, X.; Marks, R.M.

Biochem. Biophys. Res. Commun. 225, 406-412, 1996

[illegible]

```

RESULT      9
JN0473
P-selectin precursor - bovine
N:Alternate names: granule membrane protein-140
C:Species: Bos primigenius taurus (cattle)
C:date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jan-2000
C:Accession: JN0473
R:Strubel, N.A.; Meuyen, M.; Kansas, G.S.; Tedder, T.F.; Bischoff, J.
Biochem. Biophys. Res. Commun. 192, 338-344, 1993
A:title: Isolation and characterization of a bovine cDNA encoding a functional homola
A:Reference number: JN0473; MOID:93249394
A:Accession: JN0473
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-646 <STR>
A:Cross-references: GB:I12041; NID:g304246; PIDN:AAA30743.1; PID:g304247
C:Comment: This protein is a Ca2+ dependent receptor for myeloid cells.
C:Keywords: cell adhesion; glycoprotein; phosphatidyl; phosphohistidine; phosphoprotein; transmembr
F:1-41/Domaln: signal sequence #status predicted <SIG>
F:42-646/Product: P-selectin #status predicted <MAP>
F:163-194/Domaln: EGF homology <EGF>
F:200-257/Domaln: complement factor H repeat homology <FH>
F:262-319/Domaln: complement factor H repeat homology <FH2>
F:324-381/Domaln: complement factor H repeat homology <FH3>
F:386-443/Domaln: complement factor H repeat homology <FH4>
F:458-515/Domaln: complement factor H repeat homology <FH5>
F:520-577/Domaln: complement factor H repeat homology <FH6>
F:588-611/Domaln: transmembrane #status predicted <TM>
F:612-646/Domaln: intracellular #status predicted <CYT>

Query Match          67.7%; Score 451; DB 2; Length 646;
Best Local Similarity 66.7%; Pred. No. 1.le-36;
Matches 78; Conservative 10; Mismatches 29; Indels 0; Gaps 0;

OY      1 WTYHYSEKPMWNRARRCRDNTDIVAIQNKAEILEKTLPSSRSRYWTGTRKGIGIW 60
        |||||..-.-.:||:|||||||..||:|.:.:|||||||..|
Db      42 WTYYHSKTKTYMNSRAFCQYITDVAIQNKELIATLETFPIYNYSYWGIRKINMKV 101
        |||||..-.-.:||:|||||||..||:|.:.:|||||||..|

OY      61 TWVGTSKLTEEAENWGDGPNNKKNEDECVETIKRNKNADGWNDACFKLKAALC 117
        |||||..-.-.:||:|||||||..||:|.:.:|||||||..|
Db      102 TWVGTKTKLTLEEAENWADNEPNPKRNNOQCEVTIKSLASAPGNKDPCMKRRRALC 158
        |||||..-.-.:||:|||||||..||:|.:.:|||||||..|

RESULT     10
E-selectin - bovine
C:Species: Bos primigenius taurus (cattle)

```

```

Matches 77 Conservative 11 Mismatches 29 Indels 0 Gaps 0:

Qy 1 WTHYSEKPMNQARRCRDNYTDLVAIONKAEIYELEKTLPSFSRYWIGIRKIGIM 60
      ||||| : : : : : ||||| ||||| : ||| : ||||| |||||
Db 42 WTHYSTKAYSNIRKCYQNNRYTDLVAIONKNEIDILNKVLPYSSDYWKIRKNNKT 101
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||

Qy 61 TWVGTKNSTLEAEWNGDGEPPNNKKKEDCVETIYIKRNKDKAGKWNDDCHKLKALC 117
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 102 TWVGTKKALTNEAEWMAADNEPPNNKRNNEEDCVETIYIKSPAPCKWNDECHKLKHALC 158
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||

RESULT 12
153821
P-selectin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 19-May-2000
A:Accession: I53821
R:Auchampach, J.A.; Oliver, M.G.; Anderson, D.C.; Manning, A.M.
Gene 145, 231-255, 1994
A:Title: Cloning, sequence comparison and in vivo expression of the gene encoding rat
A:Reference number: I53821; MUID:94333817
A:Accession: I53821
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-768 <RES>
A:Cross-references: GB:I23088; NID:g349552; PIDN:AAA60325.1; PID:g349553
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement fac
F:32-158/Domain: C-type lectin homology <LCH>
F:163-194/Domain: EGF homology <EGF>
F:262-319/Domain: complement factor H repeat homology <FHN>
F:510-567/Domain: complement factor H repeat homology <FHO6>
F:580-637/Domain: complement factor H repeat homology <FHO7>
F:642-699/Domain: complement factor H repeat homology <FHO8>

Query Match 66.4%; Score 442; DB 2; Length 768;
Best Local Similarity 64.1%; Pred. No. 1e-35;
Matches 75; Conservative 13; Mismatches 29; Indels 0; Gaps 0;

Qy 1 WTHYSEKPMNQARRCRDNYTDLVAIONKAEIYELEKTLPSFSRYWIGIRKIGIM 60
      ||||| : : : : : ||||| ||||| : ||| : ||||| |||||
Db 42 WTHYSTKAYSNIRKCYQNNRYTDLVAIONKNEIDILNKVLPYSSDYWKIRKNNKT 101
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||

Qy 61 TWVGTKNSTLEAEWNGDGEPPNNKKKEDCVETIYIKRNKDKAGKWNDDCHKLKALC 117
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 102 TWVGTKKALTNEAEWMAADNEPPNNKRNNEEDCVETIYIKSPAPCKWNDECHKLKHALC 158
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||

RESULT 13
146709
endothelial leukocyte adhesion molecule 1 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 19-May-2000
A:Accession: I46709; I46708
R:Riariigen, J.D.; Tsang, T.C.; Rumberger, J.M.; Burns, D.K.
DNA Cell Biol. 11, 149-162, 1992
A:Title: Characterization of cDNA and genomic sequences encoding rabbit ELAM-1: Conse
A:Reference number: I46708; MUID:92189729
A:Accession: I46709
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-551 <LAR1>
A:Cross-references: GB:M91005; NID:g165006; PIDN:AAA31244.1; PID:g165007
A:Accession: I46708
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-307/'T', 309-327/'T', 329-490/'A', 492-551 <LAR2>
A:Cross-references: GB:M91004; NID:g165004; PIDN:AAA31243.1; PID:g165005
C:Genetics: ELAM1
A:Map position: 1922-q25
A:Introns: 13/1; 143/1; 179/1; 241/1; 303/1; 366/1; 429/1; 488/1; 525/1; 533/2
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement fac

```

C:Keywords: duplication; glycoprotein; tandem repeat
F:14-140/Domain: C-type lectin homology <LCH>
F:182-239/Domain: complement factor H repeat homology <FHL>
F:244-301/Domain: complement factor H repeat homology <FHL>
F:306-364/Domain: complement factor H repeat homology <FHL>
F:369-427/Domain: complement factor H repeat homology <FHL>
F:433-466/Domain: complement factor H repeat homology <FHL>
F:32,45,201,314,321,466/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 65.2%; Score 434; DB 2; Length 551;
Best Local Similarity 63.2%; Pred. No. 4.3e-35;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

Oy 1 WYHYSEKPMNWRARFCRDNYTDVAIONKAELEYLEKTLFSPRSYYWIGIRKIGIW 60
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 24 WYHESANMTYDEASACQOQRYTHLVAIONKEEIDYINSLDPSYIYWIGIRKIVNW 83
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Oy 61 TWVGTKSLTEAEENWGDEPNKKKEDCVETIYIKRKDKAGKWNDDACHTKLAALC 117
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 84 IWVGTHKPLTEGAKNMAPGEPPNNKONNEDCVETIYIKRKDKGMNDDERCSKKLAALC 140
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

RESULT 14
JC5092

E-selectin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 19-May-2000
C:Accession: JC5092
R:Winkler, H.; Brostjan, C.; Cizmadi, V.; Natarajan, G.; Anrather, J.; Bach, F.H.
Gene 176, 67-72, 1996
A:Title: The intron-exon structure of the porcine E-selectin-encoding gene.
A:Reference number: JC5092; MUID:97075911
A:Contents: endothelial cells
A:Accession: JC5092
A:Molecule type: DNA
A:Residues: 1-482 <MIN>
A:Cross-references: GB:U37521; NID:g1052974; PIDN:AA048660.1; PID:g1052975
C:Comment: This protein is a member of the selectin family of adhesion molecules.
C:Genetics: 13/1: 1421; 178/1: 237/1; 300/1: 363/1; 422/1: 459/1; 466/1
A:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor
F:13-139/Domain: C-type lectin homology <LCH>
F:181-235/Domain: complement factor H repeat homology <FHL>
F:240-298/Domain: complement factor H repeat homology <FHL>
F:303-361/Domain: complement factor H repeat homology <FHL>
F:366-420/Domain: complement factor H repeat homology <FHL>

Query Match 63.8%; Score 425; DB 2; Length 482;
Best Local Similarity 63.2%; Pred. No. 2.9e-34;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

Oy 1 WYHYSEKPMNWRARFCRDNYTDVAIONKAELEYLEKTLFSPRSYYWIGIRKIGIW 60
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 23 WYSASTEETWTFDASAYCOQRYTHLVAIONHAEIYELINSTFYASYYWIGIRKINGTW 82
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Oy 61 TWVGTKSLTEAEENWGDEPNKKKEDCVETIYIKRKDKAGKWNDDACHTKLAALC 117
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 83 TWVGTKSLTEAEENWGDEPNKKKEDCVETIYIKRKDKAGKWNDDACHTKLAALC 139
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

RESULT 15

A35046
E-selectin precursor - human
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1); LECAM-2
C:Species: Homo sapiens (man)
C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 21-Jul-2000
C:Accession: A38615; A35046; A32606
R:Collins, T.; Williams, A.; Johnston, G.I.; Kim, J.; Eddy, R.; Shows, T.; Gimbrone Jr., J.
Biol. Chem. 266, 2466-2473, 1991
A:Title: Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule 1 (ELAM-1).
A:Reference number: A38615; MUID:91115870

A:Accession: A38615
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-610 <COL>
A:Cross-references: GB:M61893; GB:M58017; NID:g182043; PIDN:AAA52377.1; PID:g182046
R:Hession, C.; Osborn, L.; Goff, D.; Chi-Rosso, G.; Vassallo, C.; Pasek, M.; Pittack, P.
Proc. Natl. Acad. Sci. U.S.A. 87, 1673-1677, 1990
A:Title: Endothelial leukocyte adhesion molecule 1: direct expression cloning and function

A:Reference number: A35046; MUID:90175359
A:Accession: A35046
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-610 <HES>

A:Cross-references: GB:M30640; NID:g182047; PIDN:AAA52377.1; PID:g182048
R:Bevilacqua, M.P.; Stengelin, S.; Gimbrone Jr., M.A.; Seed, B.
Science 243, 1160-1165, 1989

A:Title: Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophil

A:Reference number: A32606; MUID:89162047

A:Accession: A32606

A:Molecule type: mRNA

A:Residues: 1-467, 'Y', 469-610 <BEV>

A:Cross-references: GB:M24736; NID:g537523; PIDN:AAA52376.1; PID:g537524

C:Genetics:

A:Gene: GDB:SELE; ELAM; ESEL; ELAM1

A:Cross-references: GDB:120612; OMIM:131210

A:Map position: 1q22-1q25

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor

F:1-21/Domain: duplication; glycoprotein; tandem repeat; transmembrane protein

F:12-138/Domain: signal sequence #status predicted <SIG>

F:22-610/Domain: C-type lectin homology <LCH>

F:143-174/Domain: EGF homology <EGF>

F:180-237/Domain: complement factor H repeat homology <FHL>

F:242-299/Domain: complement factor H repeat homology <FHL>

F:304-362/Domain: complement factor H repeat homology <FHL>

F:367-425/Domain: complement factor H repeat homology <FHL>

F:430-488/Domain: complement factor H repeat homology <FHL>

F:493-547/Domain: complement factor H repeat homology <FHL>

F:557-578/Domain: transmembrane #status predicted <TM>

F:25,145,160,179,199,203,265,312,332,503,527/Binding site: carbohydrate (Asn) (covalent)

Query Match 62.3%; Score 415; DB 2; Length 610;
Best Local Similarity 60.7%; Pred. No. 3.6e-33;
Matches 71; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

Oy 1 WYHYSEKPMNWRARFCRDNYTDVAIONKAELEYLEKTLFSPRSYYWIGIRKIGIW 60
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 22 WSYNSTEAMTYDEASAYCOQRYTHLVAIONKEEIDYINSLDPSYIYWIGIRKIVNW 81
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Oy 61 TWVGTKSLTEAEENWGDEPNKKKEDCVETIYIKRKDKAGKWNDDACHTKLAALC 117
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 82 WVGTKPLTEGAKNMAPGEPPNNKONNEDCVETIYIKRKDKGMNDDERCSKKLAALC 138
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

RESULT 16

B42755
E-selectin precursor - mouse
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-1993 #sequence_revision 31-Dec-2000 #text_change 31-Dec-2000
C:Accession: S23174; B42755
R:Becker-Andre, M.; Van Hujsduijn, R.H.; Losberger, C.; Whelan, J.; Delamarier, J.
Eur. J. Biochem. 206, 401-411, 1992
A:Title: Murine endothelial leukocyte adhesion molecule 1 is a close structural and functional
A:Reference number: S23174; MUID:92283265
A:Accession: S23174
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <BEC>
A:Cross-references: GB:M80778; NID:g193014; PIDN:AAA37547.1; PID:g193015
J. Biol. Chem. 267, 15176-15183, 1992

QY 113 KALC 117

C1: Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 15-Sep-2000
C1: Accession: A26067; S03279; S39443; A26164; A26589; A31924; JLO132; S29107

Cell 47, 657-665, 1986
A:Title: Molecular structure of human lymphocyte receptor for immunoglobulin E.
A:Reference number: A26067; MUID:87051737
A:Accession: A26067
A:Molecule type: mRNA
A:Residues: 1-321 <KIK>
A:Cross-references: GB:M4766; NID:9182449; PIDN:AAA52435.1; PID:9182450
A:Experimental source: EBV-transformed B lymphoblastoid cells RPMI-8866
R:Suter, U.; Bastos, R.; Hofstetter, H.
Nucleic Acids Res. 15, 7295-7308, 1987
A:Title: Molecular structure of the gene and the 5'-flanking region of the human lymphoc
A:Reference number: S03279; MUID:88015596
A:Accession: S03279
A:Molecule type: DNA
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Residues: 157-284 <SUT>
A:Cross-references: GB:X06049; NID:g31316
A:Note: all exon sequences were determined but the complete sequence is not shown
R:Matsui, M.; Nunez, R.; Sachl, Y.; Lynch, R.G.; Yodoi, J.
FEBS Lett. 335, 51-56, 1993
A:Title: Alternative transcripts of the human CD23/Fc-epsilon-RII. A possible novel mech
A:Reference number: S39442; MUID:94063078
A:Accession: S39442
A:Molecule type: DNA
A:Residues: 1-7, 'D', 47-50 <MAS1>
A:Experimental source: splice form a'
A:Accession: S39443
A:Molecule type: DNA
A:Residues: 'MNPPSQ', 47-50 <MAS2>
A:Experimental source: splice form b'
R:Ludin, C.; Hofstetter, H.; Sarfati, M.; Levy, C.A.; Suter, U.; Alaimo, D.; Kilchherr,
EMBO J. 6, 109-114, 1987
A:Title: Cloning and expression of the cDNA coding for a human lymphocyte IgE receptor.
A:Reference number: A26164; MUID:87218454
A:Accession: A26164
A:Molecule type: mRNA
A:Residues: 1-268, 'T', 270-321 <LUD>
A:Cross-references: GB:X04772; NID:934002; PIDN:CAA28465.1; PID:g34003
A:Note: the codon given for 269-Asn (ACC) is inconsistent with the authors' translation
A:Note: part of this sequence, including the amino end of soluble forms of the protein,
R:Ikuta, K.; Takami, M.; Kim, C.W.; Honjo, T.; Miyoshi, T.; Tagaya, Y.; Kawabe, T.; Yodo
Proc. Natl. Acad. Sci. U.S.A. 84, 819-823, 1987
A:Title: Human lymphocyte Fc receptor for IgE: sequence homology of its cloned cDNA with
A:Reference number: A26589; MUID:87118255
A:Accession: A26589
A:Molecule type: mRNA
A:Residues: 1-321 <IKU>
A:Cross-references: GB:M15059; NID:9182447; PIDN:AAA52434.1; PID:9182448
A:Note: part of this sequence, including the amino end of soluble forms of the protein,
R:Yokota, A.; Kikutani, H.; Tanaka, T.; Sato, R.; Barsumian, E.L.; Suenura, M.; Kishimoto
Cell 55, 611-618, 1988
A:Title: Two species of human Fc-epsilon receptor II (Fc-epsilon-RII/CD23): tissue-speci
A:Reference number: A31924; MUID:89028672
A:Accession: A31924
A:Molecule type: mRNA
A:Residues: 'MNPPSQ', 8-14 <YOK>
A:Cross-references: GB:M23562; NID:9182444
A:Experimental source: splice form IIB
R:Ueteller, M.; Sarfati, M.; Delaesspe, G.
Mol. Immunol. 26, 1105-1112, 1989
A:Title: Mechanisms of formation of IgE-binding factors (soluble CD23)-I. Fc epsilon R I
A:Reference number: J10132; MUID:90220658
A:Accession: J10132
A:Molecule type: Protein
A:Residues: 1-321 <LET>
A:Experimental source: Lymphoblastoid B cell line
R:Rose, K.; Turcatti, G.; Graber, P.; Pochon, S.; Regamey, P.O.; Jansen, K.U.; Magnenat,
Biochem. J. 286, 819-824, 1992
A:Title: Partial characterization of natural and recombinant human soluble CD23.
A:Reference number: S29107; MUID:93038513
A:Accession: S29107
A:Molecule type: Protein
A:Residues: 152-166; 173-179; 189-212; 230-263; 268-306 <ROS>

R:Padlan, E.A.; Helm, B.A.
submitted to the Brookhaven Protein Data Bank, June 1993
A:Reference number: A51791; PDB:IHLI
A:Contents: annotation; conformation by theoretical model, residues 173-285
R:Baigorath, J.
submitted to the Brookhaven Protein Data Bank, November 1995
A:Reference number: A65963; PDB:IKIE
A:Contents: annotation; conformation by theoretical model, residues 173-285
C:Comment: The sequence of the splice form a is shown.
C:Comment: This receptor for the Fc portion of IgE is expressed in various hematopoie
els.
C:Comment: Splice form a is expressed constitutively in B-cells; b is expressed in ot
C:Comment: Soluble IgE-binding factors are produced by proteolytic cleavage of IgE Fc
C:Genetics:
A:Gene: GDB:FCER2; FCE2
A:Cross-references: GDB:118888; OMIM:151445
A:Map position: 19p13.3-19p13.3
A:Introns: 8/1; 46/1; 64/1; 85/1; 106/1; 127/1; 157/1; 207/3; 243/2
C:Superfamily: IgE receptor II; C-type lectin homology
C:Keywords: alternative splicing; B-cell; glycoprotein; immunoglobulin receptor; macr
F:'MNPPSQ', 47-321/Product: IgE Fc receptor II, splice form b' #status predicted <SFB>
F:1-321/Product: IgE Fc receptor II, splice form a' #status predicted <SPA>
F:1-23/Domain: intracellular #status predicted <INT>
F:1-7, 'D', 47-321/Product: IgE Fc receptor II, splice form a' #status predicted <SPAI>
F:14-20/Region: stop-transfer sequence
F:22-45/Domain: transmembrane #status predicted <TM>
F:46-321/Domain: extracellular #status predicted <EXT>
F:64-84/Region: 21-residue repeat
F:81-321/Product: soluble IgE-binding factor (37K) #status predicted <IGE>
F:85-105/Region: 21-residue repeat
F:102-321/Product: soluble IgE-binding factor (33K) #status predicted <IGB>
F:106-126/Region: 21-residue repeat
F:125-321/Product: soluble IgE-binding factor (29K) #status predicted <IGI>
F:148-321/Product: soluble IgE-binding factor (25-27K), long form #status experimenta
F:150-321/Product: soluble IgE-binding factor (25-27K), short form #status experiment
F:63/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:147-148/Cleavage site: Lys-Ieu (unidentified proteinase) #status experimental
F:149-150/Cleavage site: Arg-Met (unidentified proteinase) #status experimental
F:191-282,295-273/Distulfide bonds: #status experimental

Query Match 23.2%; Score 154.5; DB 1; Length 321;
Best Local Similarity 32.2%; Pred. No. 1e-07;
Matches 38; Conservative 32; Mismatches 47; Indels 13; Gaps 5;

OY 3 YHSEKPMWQRRRCNDNYDLVALQKAEIYLEKTPFSRSYWGIRKIG-GIW 60
DB 175 YWFGKTKQWVHARYACDMEQVLVSHSPEDDFLTRHASTGS--WIGLRNIDLKGEF 232
OY 61 YWFGTKKSLTEPEENWGDGEPNNKKNEKCEVCIYIKRKKDKGKXWDDC-HIKRAALC 117
DB 233 INVGDGH--VDYSNMAPGEPTSRSGEDCVMM-----RSGRWMDAPCDKRLGAWVC 282

RESULT 20
JX0209
lectin, galactose/N-acetylgalactosamine-specific - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: JX0209; PX0009
R:Sato, M.; Kawakami, K.; Osawa, T.; Toyoshima, S.
J. Biochem. 111, 331-336, 1992
A:Title: Molecular cloning and expression of cDNA encoding a galactose/N-acetylgalact
A:Reference number: JX0209; MUID:92268032
A:Accession: JX0209
A:Molecule type: mRNA
A:Residues: 1-304 <SAT>
A:Cross-references: GB:S36676; NID:9249360; PIDN:AMB22171.1; PID:g249361
R:Oda, S.; Sato, M.; Toyoshima, S.; Osawa, T.
J. Biochem. 104, 600-605, 1988
A:Title: Purification and characterization of a lectin-like molecule specific for gal


```

A:Reference number: PX0009; MUID:89197865
A:Molecule type: protein
A:Residues: 102-120;137,'X',139-151 <ODA>
C:Superfamily: hepatic lectin; C-type lectin homology
C:Keywords: glycoprotein; lectin; macrophage; transmembrane protein
F:36-61/Domain: transmembrane status predicted <TRA>
F:173-226/Domain: C-type lectin homology <LCH>
F:74,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      23.0%; Score 153.5; DB 2; Length 304;
Best Local Similarity 29.5%; Pred. No. 1.2e-07;
Matches 38; Conservative 18; Mismatches 52; Indels 21; Gaps 6;

OY 1 WTYH-----YSEKPMNQARARPCRDNYTDLVAIQNAEIEYLEKTLPFSRYWIGIR 54
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 WTEHGSCYWFSESESKSPREADKCYRLNSHLVYVNSLEQNFQNLRL--ANVYSWILTL 234

OY 55 KIGGIWTVV-GTNSSLTEAEAKWGDGEPNN-----KKKEDCVETIYIRKNKDAGKWNDA 108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 235 DONGFWRVWVDGTD--FEKGFKNMAPLPQDNNFQHGLGGEGECDAHT-----TTGGPMWDDV 287

OY 109 CHKLKALC 117
DB 288 CQRTFRWIC 296

RESULT 21
LNHDA
asialoglycoprotein receptor H2a - human
N:Alternate names: hepatic lectin H2a
N:Contans: HSCGP; asialoglycoprotein receptor H2a; asialoglycoprotein receptor H2b
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 10-Dec-1999
C:Accession: A25179; A39100; B39100; I37995; A49466; B49466; S14525
R:Spies, M.; Lodish, H.F.
Proc. Natl. Acad. Sci. U.S.A. 82, 6465-6469, 1985
A:Title: Sequence of a second human asialoglycoprotein receptor: conservation of two rec
A:Reference number: A25179; MUID:86016723
A:Accession: A25179
A:Molecule type: mRNA
A:Residues: 1-311 <SP1>
A:Cross-references: GB:M11025; NID:g179080; PIDN:AA859519.1; PID:g179081
R:Leiderkramer, G.Z.; Lodish, H.F.
J. Biol. Chem. 266, 1237-1244, 1991
A:Title: An alternatively spliced minlexon alters the subcellular fate of the human asial
A:Reference number: A39100; MUID:91093236
A:Accession: A39100
A:Molecule type: DNA; mRNA
A:Residues: 69-99 <LED>
A:Cross-references: GB:M38420; NID:g184395
A:Accession: B39100
A:Molecule type: DNA; mRNA
A:Residues: 69-81,87-99 <LE2>
A:Cross-references: GB:M38420; NID:g184395
R:Paletta, E.; Stockert, R.J.; Racevskis, J.
Hepatology 15, 395-402, 1992
A:Title: Differences in the abundance of variably spliced transcripts for the second asial
A:Reference number: I37995; MUID:92184202
A:Accession: I37995
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-23,43-81,87-311 <PAT>
A:Cross-references: EMBL:X55283; NID:g34354; PIDN:CAA8997.1; PID:g34355
R:Yuk, M.H.; Lodish, H.F.
J. Cell Biol. 123, 1735-1749, 1993
A:Title: Two pathways for the degradation of the H2 subunit of the asialoglycoprotein rec
A:Reference number: A49466; MUID:94103329
A:Accession: A49466
A:Molecule type: protein
A:Residues: 78-98 <YUK>

```

[illegible]

C:Keywords: acetylated amino end; glycoprotein; lectin; transmembrane protein
 F:1-23/Domain: Intracellular #status predicted <INT>
 F:24-47/Domain: transmembrane #status predicted <TRA>
 F:48-207/Domain: extracellular #status predicted <EXT>
 F:78-201/Domain: C-type lectin homology <LCH>
 F:1/Modified site: acetylated amino end (Met) #status experimental
 F:67/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 22.4%; Score 149; DB 1; Length 207;
 Best Local Similarity 30.4%; Pred. No. 2.2e-07;
 Matches 34; Conservative 21; Mismatches 41; Indels 16; Gaps 5;

QY 3 YHSEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGIR--RKIGSIW 60
 Db 93 YFSLSRSMWKAKECEHSHLIIDSYAKONFVMEFRTNER--FWIGLTDENEGEW 150
 OY 61 TWV--GTNKSILTEAEKMWGDEFPNN-KKNKEDCEVEIYIKRNKDAGKWNDDACHKLAAL 109
 Db 151 QWVGGTITRSFT----FWKGEFPNNKGFEDCAHW----TSGQWMDVYC 193

RESULT 23
 T42710
 mannose receptor, macrophage - mouse
 N:Alternate names: lambda lectin; phospholipase A2 receptor
 C:Species: Mus musculus (house mouse)
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
 C:Accession: T42710
 R:Wu, K.; Yuan, J.; Lasky, L.A.
 J. Biol. Chem. 271, 21323-21330, 1996
 A:Title: Characterization of a novel member of the macrophage mannose receptor type C 1e
 A:Reference number: Z22235; MUID:96353501
 A:Accession: T42710
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1479 <MW>
 A:Cross-references: EMBL:056734; NID:g1336073; PID:g1336074; PIDN:AAC52729.1
 C:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
 C:Keywords: membrane protein; receptor
 F:186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 22.2%; Score 148; DB 2; Length 1479;
 Best Local Similarity 30.3%; Pred. No. 2.3e-06;
 Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

QY 3 YHSEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGIR--KITGSIW 60
 Db 393 YRLQAEKRSWQESKRAOCLRGSDLLSIHMAELFFITKQIKQVEELWIGLNDKLOMNF 452
 OY 61 TWV--GTNKSILTEAEKMWGDEFPNN-KKNKEDCEVEIYIKRNKDAGKWNDDACHKLAAL 117
 Db 453 EWSGDSIVSFT----HMHFEPNPNFRDSDLCVTIW----GPEGRWNSDPCNQLSISIC 503

RESULT 24
 A48925
 mannose receptor precursor, macrophage - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A48925; S21320; PC2245
 R:Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.
 Blood 80, 2363-2373, 1992
 A:Title: Characterization of the murine macrophage mannose receptor: demonstration that
 A:Reference number: A48925; MUID:93043353
 A:Accession: A48925
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1455 <HAR>
 A:Experimental source: peritoneal macrophage (NCBIP:118733)
 A>Note: sequence extracted from NCBIP Backbone
 R:Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.

submitted to the EMBL Data Library, April 1992
 A:Description: Characterization of the murine macrophage mannose receptor: Demonstrat
 on.

A:Reference number: S21320
 A:Accession: S21320
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-302; 'N', 303-1117, 'E', 1119-1455 <HA2>
 A:Cross-references: EMBL:Z11974; NID:g52997; PIDN:CAA78028.1; PID:g52998
 R:Harris, N.; Peters, L.L.; Elcher, E.M.; Rits, M.; Raspberry, D.; Elchbaum, Q.G.; Su
 Biochem. Biophys. Res. Commun. 198, 682-692, 1994
 A:Title: The exon-intron structure and chromosomal localization of the mouse macropha
 A:Reference number: PC2245; MUID:94128116
 A:Accession: PC2245
 A:Molecule type: mRNA
 A:Residues: 35-105 <HA3>

C:Genetics:
 A:Gene: Mrc1
 A:Map position: 2
 A:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II
 C:Keywords: membrane protein; receptor
 F:168-209/Domain: fibronectin type II repeat homology <2FR>
 F:361-485/Domain: C-type lectin homology <LCH1>
 F:943-1077/Domain: C-type lectin homology <LCH2>

Query Match 21.6%; Score 144; DB 1; Length 1455;
 Best Local Similarity 28.1%; Pred. No. 5.7e-06;
 Matches 34; Conservative 22; Mismatches 51; Indels 14; Gaps 4;

QY 6 SEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGIRKIGS----IWT 61
 Db 962 NEKKRSWQARQACGLKGNLVSIENAGQAFVYHMRSSTFNATGNDIAHEMFMT 1021
 OY 62 WGTNKSILTEAEKMWGDEFPNNKKK-----EDCEVEIYIKRNKDAGKWNDDACHKLAAL 116
 Db 1022 -AGQGVHYT-----MWGKGYPGGRSSLSYEDACVVVIGNSREAGTWMDDTCSKQGYI 1076
 OY 117 C 117
 Db 1077 C 1077

RESULT 25
 S13165
 asialoglycoprotein receptor - mouse
 N:Alternate names: hepatic lectin
 C:Species: Mus musculus (house mouse)
 C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
 C:Accession: S13165
 R:Sanford, J.P.; Doyle, D.
 Biochim. Biophys. Acta 1087, 259-261, 1990
 A:Title: Mouse asialoglycoprotein receptor cDNA sequence: conservation of receptor ge
 A:Reference number: S13165; MUID:91027942
 A:Accession: S13165
 A:Molecule type: mRNA
 A:Residues: 1-301 <SAN>
 A:Cross-references: EMBL:X53042; NID:g53104; PIDN:CAA37211.1; PID:g53105
 C:Superfamily: hepatic lectin; C-type lectin homology
 C:Keywords: glycoprotein; liver; transmembrane protein
 F:170-293/Domain: C-type lectin homology <LCH>

Query Match 21.5%; Score 143.5; DB 2; Length 301;
 Best Local Similarity 28.1%; Pred. No. 1.2e-06;
 Matches 34; Conservative 22; Mismatches 50; Indels 15; Gaps 5;

QY 3 YHSEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGIRKIGSIW 62
 Db 182 YWFSRDGLTWAEDQYCOLENAHLLVINSREDFVVKHR--SQFHIMIGLTDRODSKMW 239
 OY 63 V-GTNKSILTEAEKMWGDEFPNN-----KKNKEDCEVEIYIKRNKDAGKWNDDACHKLAAL 116

Db 240 VDGTD-YRSNYRNAFTQPDNMWGHGEGEDCAEIL-----SDGHWNDFCCQVNRWV 292
Oy 117 C 117
Db 293 C 293

RESULT 26
S34198
I:GE Fc receptor II, low-affinity - rat
N:Alternate names: CD23; lymphocyte IgE receptor
C:Species: Rattus norvegicus (Norway rat)
C:Dates: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 22-Jun-1999
C:Accession: S34198
R:Flores-Romo, L.; Sheld, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayat
submitted to the EMBL Data Library, June 1993
A:Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C
A:Reference number: S34198
A:Accession: S34198
A:Molecule type: mRNA
A:Residues: 1-309 <FIO>
A:Cross-references: EMBL:X73579; NID:g313672; PIDN:CAA51981.1; PID:g313673
C:Superfamily: IgE receptor II; C-type lectin homology
C:Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr
F:1-25/Domain: intracellular #status predicted <INT>
F:14-22/Region: stop-transfer sequence
F:24-46/Domain: transmembrane #status predicted <TM>
F:47-309/Domain: extracellular #status predicted <EXT>
F:126-309/Product: soluble IgE-binding factor (29k) #status predicted <IGI>
F:149-309/Product: soluble IgE-binding factor (25-27k) #status predicted <BFI>
F:164-283/Domain: C-type lectin homology <LCH>
F:192-283,260-274/Disulfide bonds: #status predicted

Query Match 21.4%; Score 142.5; DB 1; Length 309;
Best Local Similarity 30.1%; Pred. No. 1.5e-06;
Matches 37; Conservative 23; Mismatches 40; Indels 23; Gaps 6;
Oy 3 YHSEKPMNMQRRFCNDNTDLVAIONKAEIYLEKTLPPSRSYWIGIRKIG-GIW 60
Db 176 YFEGGSKQWIOAKFTGCDLGRVLSHSQKEDFLMOHI--NKRDSWIGLQDLNMEGEF 233
Oy 61 TW-----VGTNKSLEAEENMGDGEPPNKKNKEDCEVEIYIKRNKDGAKMNDACHK-LKA 114
Db 234 VWDSGSPVGYG-----NMNPGEPNNGGCGEDCVMM-----KSGQWDAFCRSYIDA 280
Oy 115 ALC 117
Db 281 WVC 283
RESULT 27
LNMSER
I:GE Fc receptor, low-affinity - mouse
N:Alternate names: Blast-2; CD23; Fc-epsilon-RII; lymphocyte IgE receptor
C:Species: Mus musculus (house mouse)
C:Dates: 12-Feb-1993 #sequence_revision 28-Oct-1994 #text_change 22-Jun-1999
C:Accession: A43518; A33840
R:Goldnick, S.O.; Trounstein, M.L.; Yamashita, L.C.; Kehry, M.R.; Moore, K.W.
J. Immunol. 144, 1974-1982, 1990
A:Title: Isolation, characterization, and expression of cDNA clones encoding the mouse F
A:Reference number: A43518; MUID:90171598
A:Accession: A43518
A:Molecule type: mRNA
A:Residues: 1-331 <GOL>
A:Cross-references: GB:M4163; NID:g193242; PIDN:AAA37603.1; PID:g309227
R:Bettler, B.; Hofstetter, H.; Rao, M.; Yokoyama, W.M.; Kilchherr, F.; Conrad, D.H.
Proc. Natl. Acad. Sci. U.S.A. 86, 7566-7570, 1989
A:Title: Molecular structure and expression of the murine lymphocyte low-affinity recept
A:Reference number: A33840; MUID:90017519
A:Accession: A33840
A:Molecule type: mRNA
A:Residues: 1-331 <BET>

A:Cross-references: GB:M99371; NID:g193245; PIDN:AAA74898.1; PID:g193246; GB:M27150
C:Comment: This receptor for the Fc portion of IgE is expressed in various hematopoie
f B-cells.
C:Superfamily: IgE receptor II; C-type lectin homology
C:Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat;
F:1-25/Domain: intracellular #status predicted <INT>
F:14-22/Region: stop-transfer sequence
F:26-46/Domain: transmembrane #status predicted <TM>
F:47-331/Domain: extracellular #status predicted <EXT>
F:66-86/Region: 21-residue repeat
F:87-107/Region: 21-residue repeat
F:108-128/Region: 21-residue repeat
F:129-149/Region: 21-residue repeat
F:186-305/Domain: C-type lectin homology <LCH>
F:65,114/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 21.2%; Score 141.5; DB 1; Length 331;
Best Local Similarity 30.1%; Pred. No. 2e-06;
Matches 37; Conservative 23; Mismatches 40; Indels 23; Gaps 6;
Oy 3 YHSEKPMNMQRRFCNDNTDLVAIONKAEIYLEKTLPPSRSYWIGIRKIG-GIW 60
Db 198 YFEGGSKQWIOAKFTGCDLGRVLSHSQKEDFLMOHI--NKRDSWIGLQDLNMEGEF 255
Oy 61 TW-----VGTNKSLEAEENMGDGEPPNKKNKEDCEVEIYIKRNKDGAKMNDACHK-LKA 114
Db 256 VWDSGSPVGYG-----NMNPGEPNNGGCGEDCVMM-----KSGQWDAFCRSYIDA 302
Oy 115 ALC 117
Db 303 WVC 305
RESULT 28
LNRT2
I:hepatic lectin 2 - rat
N:Alternate names: asialoglycoprotein receptor RHL-2/3 (ASGP-R2/3)
C:Species: Rattus norvegicus (Norway rat)
C:Dates: 31-Mar-1988 #sequence_revision 09-Apr-1998 #text_change 22-Jun-1999
C:Accession: B28462; A28462; A31601; A26888; A23417
R:Halberg, D.F.; Wager, R.E.; Farrell, D.C.; Hildreth IV, J.; Quesenberry, M.S.; Loeb
J. Biol. Chem. 262, 9828-9838, 1987
A:Title: Major and minor forms of the rat liver asialoglycoprotein receptor are inde
A:Reference number: A28462; MUID:87250656
A:Accession: B28462
A:Molecule type: mRNA
A:Residues: 1-301 <HAL>
A:Cross-references: GB:J02762; NID:g205162; PIDN:AAA1522.1; PID:g205163
A:Accession: A28462
A:Molecule type: protein
A:Residues: 88-96, 'X', 98-118, 'X', 120, 129-158, 177-182, 'X', 184, 'X', 186-189, 192-290, 'C',
R:Sanford, J.P.; Elliott, R.W.; Doyle, D.
DNA 7, 721-728, 1988
A:Title: Asialoglycoprotein receptor genes are linked on chromosome 11 in the mouse.
A:Reference number: A31601; MUID:89170119
A:Accession: A31601
A:Molecule type: mRNA
A:Residues: 1-301 <SAN>
A:Cross-references: GB:X07636; NID:g57066; PIDN:CAA30476.1; PID:g57067
R:McPhaul, M.; Berg, P.
Mol. Cell. Biol. 7, 1841-1847, 1987
A:Title: Identification and characterization of cDNA clones encoding two homologous p
A:Reference number: A26888; MUID:87257885
A:Accession: A26888
A:Molecule type: mRNA
A:Residues: 1-152, 'A', 154-201, 'T', 203-259, 'C', 261-301 <MCP>
A:Cross-references: GB:M6347; NID:g20664; PIDN:AAA42038.1; PID:g206649
A>Note: the authors translated the codon GCA for residue 153 as Arg and ATT for resid
J. Biol. Chem. 259, 770-778, 1984
A:Title: Primary structure of the rat liver asialoglycoprotein receptor: structural e
A:Reference number: A25417; MUID:84111554

A:Accession: A25417
 A:Molecule type: Protein
 A:Residues: 201-259, 'C', 261-281, 'ND', 284-301 <DRI>
 C:Comment: Calcium is required for ligand binding.
 C:Superfamily: hepatic lectin; C-type lectin homology
 C:Keywords: endocytosis; glycoprotein; lectin; liver; receptor; transmembrane protein
 F:2-60/Domain: Intracellular #status predicted <INT>
 F:61-77/Domain: Transmembrane #status predicted <TRN>
 F:78-301/Domain: extracellular #status predicted <EXT>
 F:170-293/Domain: C-type lectin homology <LCH>
 F:97,119,165/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.1%; Score 140.5; DB 1; Length 301;
 Best Local Similarity 28.1%; Pred. No. 2.3e-06;
 Matches 34; Conservative 19; Mismatches 53; Indels 15; Gaps 5;

OY 3 YHSEKPMNQRARRFRCDNTDVAIONKAELEYLEKTLFPRSRYYWIGIRKIGITW 62
 DB 182 YWFSRDLTWAEADQYCOMENAHLLVINSREDEFFVVKHR--GAFHWIGITDKDGSWKV 239
 OY 63 V-GTNSLTDEAEWNGDEPN--KKNKEDCVETIYIKRNKDGKWNDDACHKKAAL 116
 DB 240 VDG--FYRNFKNMATAFGPDNMGHEGSEDCAEIL-----SDGLMNDNFCQVNRMA 292
 OY 117 C 117
 DB 293 C 293

RESULT 29
 JC7595
 scavenger receptor with C-type lectin type I - human
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C:Accession: JC7595

R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.
 Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001
 A:Title: Molecular cloning and functional characterization of a human scavenger receptor
 A:Reference number: JC7595; MUID:21092718; PMID:11162630
 A:Contents: Placenta

A:Accession: JC7595
 A:Molecule type: mRNA
 A:Residues: 1-742 <NAK>
 A:Cross-references: DDBJ:AB038518
 C:Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase receptor subfamily, is important for host defense. It forms a timer and plays a role in recognizing infective agents.
 C:Genetics:

A:Gene: srcl-I
 A:Map position: 16p11.32
 C:Keywords: coiled coil; glycoprotein; transmembrane protein
 F:1-33/Domain: cytosolic (amino-terminus) #status predicted <CYT>
 F:16-19/Region: internalization signal YKRF
 F:40-56/Domain: transmembrane #status predicted <TM>
 F:57-112/Domain: extracellular #status predicted <EXT>
 F:113-333/Domain: coiled coil #status predicted <CO>
 F:369-384/Region: serine/threonine-rich #status predicted
 F:443-589/Domain: collagen-like #status predicted <COL>
 F:607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CRD>

Query Match 20.5%; Score 136.5; DB 2; Length 742;
 Best Local Similarity 27.9%; Pred. No. 1.5e-05;
 Matches 34; Conservative 27; Mismatches 44; Indels 17; Gaps 6;

OY 3 YHSEKPMNQRARRFRCDNTDVAIONKAELEYLEKTLFPRSRYYWIGIRKIGITW 60
 DB 619 YWFSVKELEFEDAKLFCEKSSHLVINTREDOQWIKKQ--VGRSHWIGITLDSERENEM 677
 OY 61 TWV-GTNSLTDEAEWNGDEPNKK--KEDCVETIYIKRNKDGKWNDDACHKKAAL 115
 DB 678 KWLDTG---SPDYKNMKAGQPDNMGHGGPGEDCAGILY-----AGQNDPQCEDEVNMF 728

OY 116 LC 117
 DB 729 IC 730

RESULT 30
 A42230
 lectin M-ASGP-BP precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 20-Aug-1999
 C:Accession: A42230
 R:Li, M.; Kurata, H.; Itoh, N.; Yamashina, I.; Kawasaki, T.
 J. Biol. Chem. 265, 11295-11298, 1990
 A:Title: Molecular cloning and sequence analysis of cDNA encoding the macrophage lectin
 A:Reference number: A42230; MUID:90293078
 A:Accession: A42230
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-306 <IRA>
 A:Cross-references: GB:J05495; NID:g204302; PIDN:AAA41216.1; PID:g204303
 C:Superfamily: hepatic lectin; C-type lectin homology
 F:175-298/Domain: C-type lectin homology <LCH>

Query Match 20.0%; Score 133.5; DB 2; Length 306;
 Best Local Similarity 27.6%; Pred. No. 1.2e-05;
 Matches 34; Conservative 20; Mismatches 50; Indels 19; Gaps 6;

OY 3 YHSEKPMNQRARRFRCDNTDVAIONKAELEYLEKTLFPRSRYYWIGIRKIGITW 62
 DB 187 YWFSQSGKRPPEADKRYCOLENSNLVYVNSLAQNFQTHM--GSVYTWIGITDQNGPWRW 244
 OY 63 V-GTN--KSLTEAEWNGDEPN--KKNKEDCVETIYIKRNKDGKWNDDACHKKAAL 114
 DB 245 VDSITDEKFT---HWAKQDPDNWIGHGLGGEDECAHF-----TSDGKWNDDVCORPYR 295
 OY 115 ALC 117
 DB 296 WVC 298

RESULT 31
 A28166
 Kupffer cell receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 05-Nov-1999
 C:Accession: A38674; A28166
 R:Hoyle, G.W.; Hill, R.L.
 J. Biol. Chem. 266, 1850-1857, 1991
 A:Title: Structure of the gene for a carbohydrate-binding receptor unique to rat Kupffer cells
 A:Reference number: A38674; MUID:91107689
 A:Accession: A38674

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-550 <HO>
 A:Cross-references: GB:M5532; NID:g203362; PIDN:AAA40892.1; PID:g203363
 R:Hoyle, G.W.; Hill, R.L.
 J. Biol. Chem. 263, 7487-7492, 1988
 A:Title: Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor
 A:Reference number: A28166; MUID:86227939
 A:Accession: A28166

A:Molecule type: mRNA
 A:Residues: 1-550 <HO>
 A:Cross-references: GB:J03734; NID:g205050; PIDN:AAA41472.1; PID:g205051
 C:Superfamily: C-type lectin homology
 C:Keywords: transmembrane protein
 F:412-536/Domain: C-type lectin homology <LCH>

Query Match 19.8%; Score 132; DB 2; Length 550;
 Best Local Similarity 28.9%; Pred. No. 3.1e-05;
 Matches 33; Conservative 21; Mismatches 44; Indels 16; Gaps 6;

A:Accession: A42574
 A:Molecule type: mRNA
 A:Residues: 1-244 <KUG>
 A:Cross-references: GB:D11440; NID:g220585; PIDN:BAA02005.1; PID:g220586
 A:Experimental source: BALB/c, liver
 A:Note: Sequence extracted from NCBI backbone (NCBI:P110137)
 A:Note: parts of the sequence, including the amino end of the mature protein, were confi
 A:Accession: C42574
 A:Molecule type: protein
 A:Residues: 19,'X',21-28,'X',30-32:72-77,'H',79-80,'G',177-185,187-189,'H',191-198 <KU2>
 A:Note: source is serum of ICR mice; differences may be allotypic
 C:Genetics:
 A:Gene: Mbl2
 A:Introns: 59/1; 98/1; 121/1
 C:Superfamily: mannose-binding lectin; C-type lectin homology
 C:Keywords: calcium; endoplasmic reticulum; Golgi apparatus; homohexamer; hydroxyproline
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-244/Product: mannose-binding lectin C #status experimental <MAT>
 F:38-94/Region: collagen-like
 F:124-240/Domain: C-type lectin homology <LCH>
 F:29,34/Disulfide bonds: interchain #status predicted
 F:69/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 18.5%; Score 123.5; DB 1; Length 244;
 Best Local Similarity 29.6%; Pred. No. 8.8e-05;
 Matches 34; Conservative 14; Mismatches 52; Indels 15; Gaps 4;

OY 6 SEKPMNORARPCRDNYTDVAIONKAEIEYLEKTLFSSSYWIGIRKI---GGIWTW 62
 Db 138 SVKMSIDRKAKLCSEFGSVAITRMAEENSAIDKV---AKDIATLITDTRVREGSEDL 194
 OY 63 VGTNKSILTEAEWNGDEPNKKKKEDCEVEIYIKRNDAGKWDACHIKAKALC 117
 Db 195 TGNRVRYT---NMNDDEPNNTGDEDCVYL-----GCKMNDVPESDSFIAIC 240

RESULT 36
 LNR3
 Lectin BRA3-2 precursor - barnacle (Megabalanus rosa)
 C:Species: Megabalanus rosa
 C>Date: 31-Dec-1988 #sequence_revison 09-Sep-1994 #text_change 16-Jul-1999
 C:Accession: JCI504; A26094
 R:Nakamatsu, N.; Takeda, T.; Kojima, M.; Heishi, M.; Muramoto, K.; Kamiya, H.; Shiba, T.
 Gene 128, 251-259, 1993
 A:Title: Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene structure an
 A:Reference number: JCI503; MUID:93292994
 A:Accession: JCI504
 A:Molecule type: mRNA
 A:Residues: 1-162 <TAK>
 R:Muramoto, K.; Kamiya, H.
 Biochim. Biophys. Acta 874, 285-295, 1986
 A:Title: The amino-acid sequence of a lectin of the acorn barnacle Megabalanus rosa.
 A:Reference number: A26094
 A:Accession: A26094
 A:Molecule type: protein
 A:Residues: 25-162 <MTU>
 A:Note: 146 Arg was also found
 C:Comment: This three galactose-binding lectin is isolated from the coelomic fluid.
 C:Comment: This protein plays important roles in defense mechanisms and in development a
 C:Superfamily: tetralectin; C-type lectin homology
 C:Keywords: hemolymph; homotetramer; lectin
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-162/Product: lectin BRA3-2 #status experimental <MAT>
 F:26-150/Domain: C-type lectin homology <LCH>
 F:26-39,56-150,125-142/Disulfide bonds: #status experimental
 F:157/Disulfide bonds: interchain (to 160) #status experimental
 F:160/Disulfide bonds: interchain (to 157) #status experimental

Query Match 18.3%; Score 122; DB 1; Length 162;
 Best Local Similarity 30.6%; Pred. No. 7.9e-05;

Matches 37; Conservative 21; Mismatches 47; Indels 16; Gaps 8;
 OY 3 YHSEKPMNORARPCRDNYTD--LVAIONKAEIEYLEKTLFSSSYWIGIRKI--G 58
 Db 40 YMASTGVVRWMDAQLACQVHPGAVLATIOSLENAFISTV--SNNRLMIGLNDLDS 97
 OY 59 IWTWVGNKSLTEAEWNGDEPNKKKKEDCEVEIYIKRND--AGKWDACHIKAKAL 116
 Db 98 HYVW--SNGEAT--DETYWSSNPNPNMEN--ODCGVY---NYDVTGOWMDDDCNKKNFL 149
 OY 117 C 117
 Db 150 C 150

RESULT 37
 A28351
 pancreatic stone protein precursor - rat
 N:Alternate names: lithostathine
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Mar-1990 #sequence_revison 31-Mar-1990 #text_change 11-May-2000
 C:Accession: A28351; A39081; P10147; S34618
 R:Terazono, K.; Yamamoto, H.; Takasawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Oka
 J. Biol. Chem. 263, 2111-2114, 1988
 A:Title: A novel gene activated in regenerating islets.
 A:Reference number: A92704; MUID:88115343
 A:Accession: A28351
 A:Molecule type: mRNA
 A:Residues: 1-165 <TER>
 A:Cross-references: GB:M18962; NID:g206604; PIDN:AAA42028.1; PID:g206605
 R:Houquier, S.; Verdier, J.M.; Iovanna, J.; Dagorn, J.C.; Giorzi, D.
 J. Biol. Chem. 266, 786-791, 1991
 A:Title: Rat pancreatic stone protein messenger RNA. Abundant expression in mature ex
 A:Reference number: A39081; MUID:91093273
 A:Accession: A39081
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-165 <ROU>
 A:Cross-references: GB:M62930; GB:J05722; NID:g206462; PIDN:AAA1974.1; PID:g206463
 R:Ridrich, Z.; De Caro, A.M.; Guidoni, A.A.; Woudstra, M.E.; Roversy, M.
 Comp. Biochem. Physiol. B 93, 793-797, 1989
 A:Title: Characterization in rat pancreatic juice of a protein homologous to the huma
 A:Reference number: P10147; MUID:90031455
 A:Accession: P10147
 A:Molecule type: protein
 A:Residues: 22-69 <ADR>
 A:Experimental source: pancreas
 R:Dusetli, N.J.; Frigerio, J.M.; Dagorn, J.C.; Iovanna, J.L.
 Biochim. Biophys. Acta 1174, 99-102, 1993
 A:Title: Rapid PCR cloning and sequence determination of the rat lithostathine gene.
 A:Reference number: S34618; MUID:93326645
 A:Accession: S34618
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-165 <DUS>
 A:Cross-references: EMBL:L07512; NID:g393208; PIDN:AAA41533.1; PID:g393209
 C:Comment: This protein is found in pancreatic calculi of mammals. A peptide bond bet
 protein into an insoluble protein at a neutral pH of 5.5 to 7.5.
 C:Genetics:
 A:Introns: 21/1; 60/3; 106/3; 144/1
 C:Superfamily: tetralectin; C-type lectin homology
 C:Keywords: pyroglyutamic acid
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-165/Product: pancreatic stone protein #status predicted <MAT>
 F:35-161/Domain: C-type lectin homology <LCH>
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi

Query Match 18.2%; Score 121; DB 2; Length 165;
 Best Local Similarity 27.6%; Pred. No. 0.0001;
 Matches 34; Conservative 24; Mismatches 53; Indels 12; Gaps 6;

OY 1 WTHYSEKPMNORARPCRD--NYTDVAIONKAEIEYLEKTLFSS--RSYWIGIR-- 54

```
Db      45 YCYFMEDHLSWAEADLFCQNNNSGYLSVLSQAEGNFLASIKESGTTAAWMIGLDHP 104
Oy      55 KIGGIWTVWVGINKSLTEAEANMGSGEPNNKKNEKCELYIKRNNDACKWDDACHKRLA 114
Db      105 KNNRRMHWSGSLFL--YKSMWGTGYPNN-SNRGYCVS--VTSNSGYIKRMNDNSCDALIS 158
Oy      115 ALC 117
Db      159 FVC 161

RESULT 38
proteoglycan core protein precursor, cartilage - rat
A:Accession: A28452
N:Alternate names: aggrecan
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jan-2000
C:Accession: A92623; A23835; A28453; A28095; A28452
R:Doerge, K.; Sasakl, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 262, 17757-17767, 1987
A:Title: Complete primary structure of the rat cartilage proteoglycan core protein deduced
A:Reference number: A92623; MUID:88087070
A:Accession: A92623
A:Molecule type: mRNA
A:Residues: 1-2124 <DOE>
R:Doerge, K.; Sasakl, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 263, 10040a, 1988
A:Reference number: A30069
A:Contents: annotation; revision to residue 698
R:Doerge, K.; Fernandez, P.; Hassell, J.R.; Sasakl, M.; Yamada, Y.
J. Biol. Chem. 261, 8108-8111, 1986
A:Title: Partial cDNA sequence encoding a globular domain at the C terminus of the rat c
A:Reference number: A23835; MUID:86250698
A:Accession: A23835
A:Molecule type: mRNA
A:Residues: 1856-2124 <DO2>
A:Cross-references: GB:M13518; NID:g206104; PIDN:AAA41836.1; PID:g206105
R:Name: P.J.; Christner, J.E.; Baker, J.R.
J. Biol. Chem. 262, 17768-17778, 1987
A:Title: Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-term
A:Reference number: A28453; MUID:88087071
A:Accession: A28453
A:Molecule type: protein
A:Residues: 20-37, 'W', 39-60, 'E', 62-64, 'X', 66-69, 70-83, 84, 89-148, 'L', 150-238, 'S', 240, 'A',
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2124/Product: proteoglycan core protein #status predicted <MAT>
F:44-135/Domain: immunoglobulin homology <IMM>
F:170-247/Domain: link protein repeat homology <LNK1>
F:268-349/Domain: link protein repeat homology <LNK2>
F:504-581/Domain: link protein repeat homology <LNK3>
F:602-683/Domain: link protein repeat homology <LNK4>
F:1914-2034/Domain: C-type lectin homology <LCH>
F:2041-2097/Domain: complement factor H repeat homology <FHD>
F:126,229,333,387,611,667,1842/Binding site: carbohydrate (Asn) (covalent) #status predi
```

```
A:55182
aggrecan precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: A55182; S55329; S50207; S51355; I78532; I58123
R:Walcz, E.; Deak, F.; Erhardt, P.; Coulter, S.N.; Fueleop, C.; Horvath, P.; Doege, K
Genomics 22, 364-371, 1994
A:Title: Complete coding sequence, deduced primary structure, chromosomal localizatio
A:Reference number: A55182; MUID:95104847
A:Accession: A55182
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-2132 <WAL>
A:Cross-references: GB:I07049; NID:g678541; PIDN:AAC37670.1; PID:g191772
R:Watanabe, H.; Gao, L.; Sugiyama, S.; Doege, K.; Kimata, K.; Yamada, Y.
Biochem. J. 308, 433-440, 1995
A:Title: Mouse aggrecan, a large cartilage proteoglycan: protein sequence, gene struc
A:Reference number: S55329; MUID:95289972
A:Accession: S55329
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-567, 'G', 569-1728, 'I', 1730-2132 <WAT1>
A:Cross-references: GB:U22901; NID:g886014
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues
A:Reference number: S50206; MUID:95035091
A:Accession: S50207
A:Molecule type: mRNA
A:Residues: 1-567, 'G', 569-1728, 'I', 1730-2132 <WAT1>
A:Cross-references: GB:U22901; NID:g886014
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues
A:Reference number: S50206; MUID:95035091
A:Accession: S50207
A:Molecule type: mRNA
A:Residues: 350-481, 'R', 483-506 <GLU1>
A:Cross-references: EMBL:X80279; NID:g673432
R:Glumoff, V.
submitted to the EMBL Data Library, July 1994
A:Reference number: S51355
A:Accession: S51355
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 350-383, 'CPVMSOREPMAA' <GLU2>
A:Cross-references: EMBL:X80279
R:Watanabe, H.; Kimata, K.; Iino, S.; Strong, D.; Gao, L.Y.; Kozak, C.A.; Yamada, Y.
Nature Genet. 7, 154-157, 1994
A:Title: Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the agg
A:Reference number: I58123; MUID:95004579
A:Accession: I78532
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 211-326 <WAT2>
A:Cross-references: GB:S73722; NID:g765215; PIDN:AAB32160.1; PID:g765216
A:Accession: I58123
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 211-240, 'MCTASLRMRKRSRMRPQNSPSRRPPTS', 'AGCGHAMPPQASSTWGRVAVTCAALAGW'
A:Cross-references: GB:S73720; NID:g765211; PIDN:AAB32159.1; PID:g765212
C:Genetics:
A:Map position: 7
A:Insertions: 253/1
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C:Keywords: cartilage; extracellular matrix
F:1-19/Domain: signal sequence #status predicted <SIG>
F:44-135/Domain: immunoglobulin homology <IMM>
F:170-247/Domain: link protein repeat homology <LNK1>
F:268-349/Domain: link protein repeat homology <LNK2>
F:504-581/Domain: link protein repeat homology <LNK3>
F:602-683/Domain: link protein repeat homology <LNK4>
F:1922-2042/Domain: C-type lectin homology <LCH>
F:2049-2105/Domain: complement factor H repeat homology <FHD>

Query Match 18.2%; Score 121; DB 1; Length 2132;
Best Local Similarity 31.7%; Pred. No. 0.0016;
Matches 38; Conservative 17; Mismatches 45; Indels 20; Gaps 8;

Oy      1 WT-----YHSEKPMNQARRRRCRDNYTDLVAIQNAEIEYLEKTLPFSSYYWIGI- 53
```

DB 1926 WTKFGCHCYRHFDPRETVDAERRCRQSHLSSTIVPEDEEFVNKN---AODYQWIGLN 1982
 QY 54 -RKIGGIWTVGVINKSLTEAEANWGDGPNN-KKNKEDC-VEIYIKRNKAGKNDACH 110
 DB 1983 DRTIEGDFRW-SPQSHL--QFEKWRPNQPDNFRATGEDCVYMIWHER---GEWNDVPCN 2035

RESULT 40

LNRTL

hepatic lectin - rat

N:Alternate names: ASGP, asialoglycoprotein receptor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Feb-1986 #sequence_revision 04-Dec-1986 #text_change 22-Jun-1999
 C:Accession: A92497; A94020; B94020; A54727; A03166
 R:Leung, J.O.; Holland, E.C.; Drickamer, K.
 J. Biol. Chem. 260, 12523-12527, 1985
 A:Title: Characterization of the gene encoding the major rat liver asialoglycoprotein re
 A:Reference number: A92497; MUID:86008335
 A:Accession: A92497
 A:Molecule type: DNA
 A:Residues: 1-284 <LEU>

A:Cross-references: GB:K02817; NID:q206646; PIDN:AAA42037.1; PID:q206647
 R:Holland, E.C.; Leung, J.O.; Drickamer, K.
 Proc. Natl. Acad. Sci. U.S.A. 81, 7338-7342, 1984
 A:Title: Rat liver asialoglycoprotein receptor lacks a cleavable NH-2-terminal signal se
 A:Reference number: A94020; MUID:85063786
 A:Accession: A94020
 A:Molecule type: mRNA
 A:Residues: 1-60, 'R', 62-210 <HOL>
 A:Experimental source: clone 22; clone 1
 A:Accession: B94020
 A:Molecule type: mRNA
 A:Residues: 92-284 <HO2>
 A:Note: Clone 22 codes for a terminator at residue 210
 R:Watts, C.
 Biosci. Rep. 6, 527-534, 1986
 A:Title: Isolation and expression of cDNA clones for a rat liver asialoglycoprotein rece
 A:Reference number: A54727; MUID:87026895
 A:Accession: A54727

A:Molecule type: mRNA
 A:Residues: 12-284 <WAT>
 A:Cross-references: GB:M21770; NID:q202985; PIDN:AAA40764.1; PID:q202988
 A:Experimental source: liver
 C:Comment: Two types of rat hepatic lectin have been identified, RHL-1 and RHL-2/3, hav
 C:Comment: After removal of sialic acid monomers from the complex carbohydrate moieties
 C:Comment: The unusual orientation of this protein across the membrane is postulated to
 C:Genetics:
 A:Introns: 23/1; 62/1; 94/1; 118/1; 147/1; 197/3; 233/2
 C:Superfamily: hepatic lectin; C-type lectin homology
 C:Keywords: endocytosis; glycoprotein; lectin; receptor; transmembrane protein
 F:2-284/Product: hepatic lectin #status predicted <MAT>
 F:2-39/Domain: intracellular #status predicted <INT>
 F:40-60/Domain: transmembrane #status predicted <TM>
 F:61-284/Domain: extracellular #status predicted <EXT>
 F:153-276/Domain: C-type lectin homology <LCH>
 F:75,78,146/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.1%; Score 120.5; DB 1; Length 284;
 Best Local Similarity 26.2%; Pred. No. 0.0002;
 Matches 32; Conservative 20; Mismatches 53; Indels 17; Gaps 6;

QY 3 YHSEKPMNMORARRCRDNYDVAIQKAEIYELEKTL-PSRSRYWYIGIRKIGIWT 61
 DB 165 YWFSSSVKRPWTEADKYCOLENAHLVVTSMWEGRFVQOHGPIPLT--WIGLTDQNGPWK 221

QY 62 WV-GTNKSLTEAEANWGDGPNN----KKNKEDVEIYIKRNKAGKNDACHKLKAA 115
 DB 222 WYDGTD--YETGFKNWRPQOPDWDWGHGLGGECAHF-----TTDGHMNDVCRRPYRW 274

QY 116 LC 117
 :|

DB 275 VC 276

Search completed: September 7, 2002, 10:15:25
 Job time: 203 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:12:52 : Search time 85.16 Seconds
(without alignments)
347.469 Million cell updates/sec

Title: US-09-119-209-2_COPY_39_155

Perfect score: 666

Sequence: 1 WTYHYSEKPMNQRARFCR.....NKDAGKWNDDACHKLKALC 117

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 869228 seqs, 252910224 residues

Total number of hits satisfying chosen parameters: 869228

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	323	5	US-09-791-537-53485 Sequence 53485, A
2	666	100.0	341	6	US-10-211-364-1171 Sequence 1171, Ap
3	666	100.0	341	6	US-10-212-054-1328 Sequence 1328, Ap
4	666	100.0	341	6	US-10-212-778-1158 Sequence 1158, Ap
5	666	100.0	363	5	US-09-791-537-58446 Sequence 58446, A
6	666	100.0	372	5	US-09-791-537-22816 Sequence 22816, A
7	666	100.0	372	5	US-09-791-537-42657 Sequence 42657, A
8	666	100.0	372	5	US-09-791-537-69658 Sequence 69658, A
9	666	100.0	372	5	US-09-791-537-84593 Sequence 84593, A
10	666	100.0	372	5	US-09-791-537-15267 Sequence 15267, A
11	666	100.0	385	5	US-09-791-537-51391 Sequence 51391, A
12	666	100.0	385	5	US-09-791-537-53844 Sequence 53844, A
13	666	100.0	385	5	US-09-791-537-113060 Sequence 113060, A
14	666	99.4	372	5	US-09-791-537-42659 Sequence 42659, A
15	631	94.7	372	5	US-09-791-537-42655 Sequence 42655, A
16	631	94.7	372	5	US-09-791-537-42658 Sequence 42658, A
17	598	89.8	376	5	US-09-791-537-13214 Sequence 13214, A
18	581	87.2	370	5	US-09-791-537-50403 Sequence 50403, A
19	565	84.8	360	5	US-09-791-537-60503 Sequence 60503, A
20	565	84.8	372	5	US-09-791-537-37750 Sequence 37750, A
21	557	83.6	372	5	US-09-791-537-6693 Sequence 6693, Ap
22	557	83.6	372	5	US-09-791-537-81233 Sequence 81233, A
23	464	69.7	769	5	US-09-791-537-50409 Sequence 50409, A
24	455	68.3	768	5	US-09-791-537-20989 Sequence 20989, A
25	455	68.3	768	5	US-09-791-537-37753 Sequence 37753, A
26	451	67.7	646	5	US-09-791-537-84829 Sequence 84829, A

27	445	66.8	485	5	US-09-791-537-50404 Sequence 50404, A
28	445	66.8	616	5	US-09-791-537-33205 Sequence 33205, A
29	445	66.8	740	5	US-09-791-537-33023 Sequence 33023, A
30	445	66.8	740	5	US-09-791-537-32024 Sequence 32024, A
31	445	66.8	830	5	US-09-791-537-22819 Sequence 22819, A
32	445	66.8	830	5	US-09-791-537-35618 Sequence 35618, A
33	442	66.4	768	5	US-09-791-537-86038 Sequence 86038, A
34	434	65.2	551	5	US-09-791-537-44925 Sequence 44925, A
35	434	65.2	551	5	US-09-791-537-121836 Sequence 121836, A
36	434	65.2	754	5	US-09-791-537-14651 Sequence 14651, A
37	425	63.8	482	5	US-09-791-537-41672 Sequence 41672, A
38	425	63.8	482	5	US-09-791-537-11285 Sequence 11285, A
39	425	63.8	484	5	US-09-791-537-50405 Sequence 50405, A
40	420	63.1	649	5	US-09-791-537-93873 Sequence 93873, A
41	415	62.3	162	5	US-09-791-537-23269 Sequence 23269, A
42	415	62.3	196	6	US-10-143-788-916 Sequence 916, App
43	415	62.3	610	1	PCT-US02-23913-357 Sequence 357, App
44	415	62.3	610	5	US-09-791-537-22813 Sequence 22813, A
45	415	62.3	610	5	US-09-791-537-121834 Sequence 121834, A

ALIGNMENTS

```
RESULT
1
US-09-791-537-53485
: Sequence 53485, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biocomix, Inc.
: APPLICANT: Debe, Derek
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 53485
: LENGTH: 323
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-791-537-53485

Query Match      100.0%; Score 666; DB 5; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.8e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WTYHYSEKPMNQRARFCRDYVTDVATONKAETLEYEKLTPFSRYWIGIRKIGTW 60
      |||||||
DB      52 WTYHYSEKPMNQRARFCRDYVTDVATONKAETLEYEKLTPFSRYWIGIRKIGTW 111
      |||||||

QY      61 TWVGTRKSLTEAEWNGDEPNKKKEDCEVEIYIRNRKDGKWNDDACHKLKALC 117
      |||||||
DB      112 TWVGTRKSLTEAEWNGDEPNKKKEDCEVEIYIRNRKDGKWNDDACHKLKALC 168
      |||||||

RESULT
2
US-10-211-364-1171
: Sequence 1171, Application US/10211364
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P1216CIN
: CURRENT APPLICATION NUMBER: US/10/211,364
: CURRENT FILING DATE: 2002-08-05
: PRIOR APPLICATION NUMBER: 09/760,486
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: 60/179,065
: PRIOR FILING DATE: 2000-01-31
: PRIOR APPLICATION NUMBER: 60/180,628
: PRIOR FILING DATE: 2000-02-04
```

```

; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1778
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1171
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-211-364-1171
```

```

Query Match          100.0%; Score 666; DB 6; Length 341;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 60
    |||||||
Db 61 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 120

QY 61 TWVGTNKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKRLKALC 117
    |||||||
Db 121 TWVGTNKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKRLKALC 177
```

```

RESULT 3
; US-10-212-054-1328
; Sequence 1328, Application US/10212054
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P212C1N
; CURRENT APPLICATION NUMBER: US/10/212,054
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2164
; Prior application removed - See file Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1328
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-054-1328
```

```

Query Match          100.0%; Score 666; DB 6; Length 341;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 60
    |||||||
Db 61 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 120

*QY 61 TWVGTNKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKRLKALC 117
```

```

Db 121 TWVGTNKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKRLKALC 177
```

```

RESULT 4
; US-10-212-778-1158
; Sequence 1158, Application US/10212778
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM026C1N
; CURRENT APPLICATION NUMBER: US/10/212,778
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/758,449
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1478
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1158
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-778-1158
```

```

Query Match          100.0%; Score 666; DB 6; Length 341;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 60
    |||||||
Db 61 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 120

QY 61 TWVGTNKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKRLKALC 117
    |||||||
Db 121 TWVGTNKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKRLKALC 177
```

```

RESULT 5
; US-09-791-537-58446
; Sequence 58446, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomatik, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58446
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-58446
```

```

Query Match          100.0%; Score 666; DB 5; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.1e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 60
    |||||||
Db 52 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYELEKTLPSRSYYWIGIRKIGIW 111
```

Oy 61 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
Db 112 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 168

RESULT 6
US-09-791-537-22816

; Sequence 22816, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22816
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-22816

Query Match 100.0%; Score 666; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 60
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 98

Oy 61 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
Db 99 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 7
US-09-791-537-42657

; Sequence 42657, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42657
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pan troglodytes
US-09-791-537-42657

Query Match 100.0%; Score 666; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 60
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 98

Oy 61 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
Db 99 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 8
US-09-791-537-69658

; Sequence 69658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-69658

Query Match 100.0%; Score 666; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 60
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 98

Oy 61 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
Db 99 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 9
US-09-791-537-84593

; Sequence 84593, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84593
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-84593

Query Match 100.0%; Score 666; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 60
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETLEYEKLTPFSRSYYWIGIRKIGIW 98

Oy 61 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
Db 99 TWGTTNKSLTTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 10
US-09-791-537-152667

; Sequence 152667, Application US/09791537
; GENERAL INFORMATION:

```
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 152667
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-152667
```

```
Query Match          100.0%; Score 666; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 60
    |||||||
Db 39 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 98
```

```
QY 61 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 99 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 155
```

```
RESULT 11
US-09-791-537-51391
; Sequence 51391, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51391
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-51391
```

```
Query Match          100.0%; Score 666; DB 5; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,3e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 60
    |||||||
Db 52 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 111
```

```
QY 61 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 112 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 168
```

```
RESULT 12
US-09-791-537-53844
; Sequence 53844, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
```

```
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53844
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53844
```

```
Query Match          100.0%; Score 666; DB 5; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,3e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 60
    |||||||
Db 52 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 111
```

```
QY 61 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 112 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 168
```

```
RESULT 13
US-09-791-537-113060
; Sequence 113060, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113060
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-113060
```

```
Query Match          100.0%; Score 666; DB 5; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,3e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 60
    |||||||
Db 52 WTYHYSEKPMNQARARFCRDNTYDVAIONKAEIYELEKTLPSRSYIWIGIRKIGIW 111
```

```
QY 61 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 112 TWVGTKNSLTEEAENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 168
```

```
RESULT 14
US-09-791-537-42659
; Sequence 42659, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
```

```

:
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 42659
:
: LENGTH: 372
:
: TYPE: PRT
:
: ORGANISM: Pongo pygmaeus
: US-09-791-537-42659

```

Query Match	99.4%;	Score 662;	DB 5;	Length 372;
Best Local Similarity	99.1%;	Pred. No. 6e-64;		
Matches 116; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

Oy 1 WTYYHSEKPMWQRRARRRCRDNYTDLVAIONKAIEIYLEKTLPFSRSYYWGIRRIIGGIW 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 WTYHHSEKPMWQRRARRCRCRENTDLDVAIONKAIEIYLEKTLPFSRSYYIWGIRRIIGGIW 980

Qy	61	TWVGTSNSTEEAEWNGDGEPPNNKKNKEDCEVEIYIRKNDAGKYNDDACHLKALC	117
Db	99	TWVGTSNSTEEAEWNGDGEPPNNKKNKEDCEVEIYIRKNDAGKYNDDACHLKALC	155

RESULT 15
US-09-791-537-42655
; Sequence 42655, Application US/09791537

: APPLICANT: Blonomix, Inc.
 : APPLICANT: Debe, Derek
 : APPLICANT: Danzer, Joseph
 : TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
 : TITLE OF INVENTION: METHODS OF USE THEREOF
 : FILE REFERENCE: 261/210
 : CURRENT APPLICATION NUMBER: US/09/791,537
 : CURRENT FILING DATE: 2001-02-22
 : NUMBER OF SEQ ID NOS: 153055
 : SOFTWARE: PatentIn version 3.0
 : SEQ ID NO 42655
 : LENGTH: 372

ORGANISM: Macaca mulatta
US-09-791-537-42655

Query Match	94.7%;	Score 631;	DB 5;	Length 372;
Best Local Similarity	94.0%;	Pred. No. 1.4e-60;		
Matches 110; Conservative	3;	Mismatches 4;	Indels 0;	Gaps 0;

OY 1 WTYHYSEKPMNQARARECRDNVTDLVAIONKAIEIYLEKTLPFSRSYYWIGIRKGIW 60
||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 39 WTYHYSENPMQAKARECRENTDVLAIONKAIEIYLEKTLPFSPSYYWIGIRKGIW 98

QY 61 TWVGTSKSLTEAEANWGDGEPNNKKNKEDCVEIYIKRNKDAGKNDDACHKLKALC 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 99 TWWGTNKSLSLOEAEANWGDGEPNNKKNKEDCVEIYIKRKKDAGKNDDACHPKAALC 155

RESULT 16
US-09-791-537-42658
; Sequence 42658, Application US/09791537
; GENERAL INFORMATION:

APPLICANT: Decker, Decker
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
TITLE OF INVENTION: METHODS OF USE THEREOF

```

; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42658

```

```

; LENGTH: 372
; TYPE: PRT
; ORGANISM: Papio hamadryas
US-09-791-537-42658

```

Query Match	94.7%	Score 631;	DB 5;	Length 372;
Best Local Similarity	94.0%	Pred. No. 1,4e-60;		
Matches 110: Conservative	3;	Mismatches 4;	Indels 0;	Gaps 0;

```
QY      1 WTYHSEKPMNQRRARFCRDNTYTDLVAIONKAIEYLEKTLPFSRYWIGIRKIGIW   600
          ||||| :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     39 WTYYHSNPMMOKARRCRENNTDVAIONKAIEYLEKTLPFSPSYWIGIRKIGIW   988
```

[illegible]

RESULT 17
US-09-791-537-132144
; Sequence 132144, Application US/09791537

```

: APPLICANT: Danzer, Joseph
:
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
:
: TITLE OF INVENTION: METHODS OF USE THEREOF
:
: FILE REFERENCE: 261/210
:
: CURRENT APPLICATION NUMBER: US/09/791,537
:
: CURRENT FILING DATE: 2001-02-22
:
: NUMBER OF SEQ ID NOS: 153055
:
: SOFTWARE: PatentIn version 3.0
:
: SEQ ID NO 132144
:
: LENGTH: 376
:
: TYPE: PRT
:
: ORGANISM: Oryctolagus cuniculus
:
: US-09-791-537-132144

```

Query Match	89.8%	Score 598;	DB 5;	Length 376;
Best Local Similarity	88.0%;	Pred. No. 5.9e-57;		
Matches 103; Conservative	7;	Mismatches 7;	Indels 0;	Gaps 0;

QY 1 WTYHSEKPMNQARARRCDNYTDVLAIONKAEIEYLEKTLPFSRSYYWIGIRKIGIW 60
 ||||| :||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
Dd 39 WTYHSEKPMNMERARKFCRENYTDVLAIONKGEIETLEKTLDFSRSYTWIGIRKIGNIT 98

QY 61 TWVGTNKS/LTEAEANWGDGEPNNKKNKECDVEIYIKRNKDAGKKNDACHLKAALC 117
||||| : ||||| : ||||| : ||||| : |||||
Db 99 TWVGTNKS/LTAAEAWNGEGEPNNKKTKECDVEIYIKRLRDSGWINDSCQRKAALC 155

RESULT 18
US-09-791-537-50403
; Sequence 50403, Application US/09791537

INVENTOR: Schenck, Jr., Eric
 APPLICANT: Debe, Derek
 DOR: Danzer, Joseph
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
 TITLE OF INVENTION: METHODS OF USE THEREOF

```

?
?
? CURRENT FILING DATE: 2001-02-22
? CURRENT APPLICATION NUMBER: US/09/191,533
?
? NUMBER OF SEQ ID NOS: 153055
?
? SOFTWARE: PatentIn version 3.0
?
```

```

; SEQ ID NO 50403
; LENGTH: 370
; TYPE: prt
; ORGANISM: Bos taurus
US-09-791-537-50403

```

Query Match	84.8%;	Score 565;	DB 5;	Length 372;
Best Local Similarity	82.9%;	Pred. No. 2.3e-53;		
Matches 97;	Conservative	9;	Mismatches 11;	Indels 0;
				Gaps 0;

```

QY      1 WTHYSEKPNWQRRARRECRDNTDLVAIONKAELEYLEKTLPSRSRYTWIGIRIGGIW 60
          ||||| : ||| : ||||| ||||| ||||| : ||||| ||||| |
Db      39 WTHYSESRMNMWENARKCKHNYTDLVAIONKRELEYLEKTLPKNPPTYWIGIRIRIGKTW 98

```

61 TWVGTNKS L TEAE N WGDGE P N N K N K E D V E I Y I K R N K D A G K W N D D A C H K L K A A L C 117

Db 99 TWWGTKTLTEAEENMGTEGEPNNKSKEDCVELIYIKRERDSGKWNDDACHKKRAALC 155

RESULT 23

US-09-791-537-50409
; Sequence 50409, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50409
; LENGTH: 769
; TYPE: PRT
; ORGANISM: Ovis aries
US-09-791-537-50409

Query Match 69.7%; Score 464; DB 5; Length 769;
Best Local Similarity 68.4%; Pred. No. 6e-42; Indels 0; Gaps 0;

Matches 80; Conservative 10; Mismatches 27; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQARARFCRDNYTDLVAIONKAEILEKTLPEFSRSYYWIGIRKIGIWM 60

Db 42 WTYHYSKDKYNNYSRAFCQKYYTDLVAIONKNEIAYLWETIPIYNSYYWIGIRKINNM 101

QY 61 TWWGTKSLTEAEENMGDEPNKKKKEDCVELIYIKRNKDGKWNDDACHKKRAALC 117
Db 102 TWWGTKTLTEAEENMGADNEPNKKKNQDCVELIYIKSPAGPKWNDEPCGKRRKRALC 158

RESULT 24

US-09-791-537-20989
; Sequence 20989, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20989
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-20989

Query Match 68.3%; Score 455; DB 5; Length 768;
Best Local Similarity 67.5%; Pred. No. 5.7e-41; Indels 0; Gaps 0;
Matches 79; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQARARFCRDNYTDLVAIONKAEILEKTLPEFSRSYYWIGIRKIGIWM 60

Db 42 WTYHYSKDKYNNYSRAFCQKYYTDLVAIONKNEIAYLWETIPIYNSYYWIGIRKINNM 101

QY 61 TWWGTKSLTEAEENMGDEPNKKKKEDCVELIYIKRNKDGKWNDDACHKKRAALC 117
Db 102 TWWGTKTLTEAEENMGADNEPNKKKNQDCVELIYIKSPAGPKWNDEPCGKRRKRALC 158

RESULT 25

US-09-791-537-37753
; Sequence 37753, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37753
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-37753

Query Match 68.3%; Score 455; DB 5; Length 768;
Best Local Similarity 67.5%; Pred. No. 5.7e-41; Indels 0; Gaps 0;

Matches 79; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQARARFCRDNYTDLVAIONKAEILEKTLPEFSRSYYWIGIRKIGIWM 60

Db 42 WTYHYSKDKYNNYSRAFCQKYYTDLVAIONKNEIAYLWETIPIYNSYYWIGIRKINNM 101

QY 61 TWWGTKSLTEAEENMGDEPNKKKKEDCVELIYIKRNKDGKWNDDACHKKRAALC 117
Db 102 TWWGTKTLTEAEENMGADNEPNKKKNQDCVELIYIKSPAGPKWNDEPCGKRRKRALC 158

RESULT 26

US-09-791-537-84829
; Sequence 84829, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84829
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-84829

Query Match 67.7%; Score 451; DB 5; Length 646;
Best Local Similarity 66.7%; Pred. No. 1.3e-40; Indels 0; Gaps 0;
Matches 78; Conservative 10; Mismatches 29; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQARARFCRDNYTDLVAIONKAEILEKTLPEFSRSYYWIGIRKIGIWM 60

Db 42 WTYHYSKDKYNNYSRAFCQKYYTDLVAIONKNEIAYLWETIPIYNSYYWIGIRKINNM 101

QY 61 TWWGTKSLTEAEENMGDEPNKKKKEDCVELIYIKRNKDGKWNDDACHKKRAALC 117
Db 102 TWWGTKTLTEAEENMGADNEPNKKKNQDCVELIYIKSPAGPKWNDEPCGKRRKRALC 158

RESULT 27

US-09-791-537-50404
; Sequence 50404, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek

```

; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50404
; LENGTH: 485
; TYPE: PRF
; ORGANISM: Bos taurus
US-09-791-537-50404
```

```

Query Match          66.8%; Score 445; DB 5; Length 485;
Best Local Similarity 65.8%; Pred. No. 4.1e-40;
Matches 77; Conservative 14; Mismatches 26; Indels 0; Gaps 0;
```

```

QY 1 WYHYSEKPMNQARFCRDNYTDLVAIONKAEIYELEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 23 WYHASTEMMTFEARDYCOKTYTALVAIONOEIYELNSTFESYSPSYWIGIRKINGTW 82
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

QY 61 TWGCTNKSITFEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHIKKALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 83 TWIGTKKALTNEAEENWADNEPNKRNNEKDCVEIYIKRSPAGKWNDEHCKLKKHALC 139
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

RESULT 28
US-09-791-537-32025
; Sequence 32025, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32025
; LENGTH: 616
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-32025
```

```

Query Match          66.8%; Score 445; DB 5; Length 616;
Best Local Similarity 65.8%; Pred. No. 5.5e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;
```

```

QY 1 WYHYSEKPMNQARFCRDNYTDLVAIONKAEIYELEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHASTKAYSWNISRYCQNYTDLVAIONKNEIDYLNKVLPPYSYWIGIRKNNKTW 101
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

QY 61 TWGCTNKSITFEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHIKKALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGCTKALTNEAEENWADNEPNKRNNEKDCVEIYIKRSPAGKWNDEHCKLKKHALC 158
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

RESULT 29
US-09-791-537-32023
; Sequence 32023, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
```

```

; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32023
; LENGTH: 740
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-32023
```

```

Query Match          66.8%; Score 445; DB 5; Length 740;
Best Local Similarity 65.8%; Pred. No. 6.8e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;
```

```

QY 1 WYHYSEKPMNQARFCRDNYTDLVAIONKAEIYELEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHASTKAYSWNISRYCQNYTDLVAIONKNEIDYLNKVLPPYSYWIGIRKNNKTW 101
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

QY 61 TWGCTNKSITFEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHIKKALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGCTKALTNEAEENWADNEPNKRNNEKDCVEIYIKRSPAGKWNDEHCKLKKHALC 158
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

RESULT 30
US-09-791-537-32024
; Sequence 32024, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32024
; LENGTH: 740
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-32024
```

```

Query Match          66.8%; Score 445; DB 5; Length 740;
Best Local Similarity 65.8%; Pred. No. 6.8e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;
```

```

QY 1 WYHYSEKPMNQARFCRDNYTDLVAIONKAEIYELEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHASTKAYSWNISRYCQNYTDLVAIONKNEIDYLNKVLPPYSYWIGIRKNNKTW 101
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

QY 61 TWGCTNKSITFEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHIKKALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGCTKALTNEAEENWADNEPNKRNNEKDCVEIYIKRSPAGKWNDEHCKLKKHALC 158
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

```

RESULT 31
US-09-791-537-22819
; Sequence 22819, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22819
; LENGTH: 830
```


TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-22819

Query Match 66.8%; Score 445; DB 5; Length 830;
Best Local Similarity 65.8%; Pred. No. 7.8e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETIELEKTLPEFSRSYWGIRKIGIM 60
||||| : : : : : ||||||| : : : : : ||||||| :
DB 42 WTYHSTKAYSNISRKTCQNRITDLVAIONKNEIDYLNKVLPIYSSYIWIGIRKNNKW 101

OY 61 TWVGTKSLTEBAENWGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117
||||| : : : : : ||||||| : : : : : ||||||| :
DB 102 TWVGTKKALTNEAENWADNEPNKKRNNECDEVEIYIKSPAPGKWDEHCKKKHALC 158

RESULT 32
US-09-791-537-35618
Sequence 35618, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
CURRENT APPLICATION NUMBER: US/09/791,537
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 35618
LENGTH: 830
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-35618

Query Match 66.8%; Score 445; DB 5; Length 830;
Best Local Similarity 65.8%; Pred. No. 7.8e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETIELEKTLPEFSRSYWGIRKIGIM 60
||||| : : : : : ||||||| : : : : : ||||||| :
DB 42 WTYHSTKAYSNISRKTCQNRITDLVAIONKNEIDYLNKVLPIYSSYIWIGIRKNNKW 101

OY 61 TWVGTKSLTEBAENWGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117
||||| : : : : : ||||||| : : : : : ||||||| :
DB 102 TWVGTKKALTNEAENWADNEPNKKRNNECDEVEIYIKSPAPGKWDEHCKKKHALC 158

RESULT 33
US-09-791-537-86038
Sequence 86038, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
CURRENT APPLICATION NUMBER: US/09/791,537
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 86038
LENGTH: 768
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-791-537-86038

Query Match 66.4%; Score 442; DB 5; Length 768;
Best Local Similarity 64.1%; Pred. No. 1.5e-39;
Matches 75; Conservative 13; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETIELEKTLPEFSRSYWGIRKIGIM 60
||||| : : : : : ||||||| : : : : : ||||||| :
DB 42 WTYNSTKAYSNNSHAFKRFHTDLVAIONKNEIHLNDVLPYVNSYIWIGIRKINNKW 101

OY 61 TWVGTKSLTEBAENWGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117
||||| : : : : : ||||||| : : : : : ||||||| :
DB 102 TWVGTKKALTNEAENWADNEPNKKRNQDCVEIYIKSNSAPGKWDEPCKKRRALC 158

RESULT 34
US-09-791-537-44925
Sequence 44925, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
CURRENT APPLICATION NUMBER: US/09/791,537
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 44925
LENGTH: 551
TYPE: PRT
ORGANISM: Oryctolagus cuniculus
US-09-791-537-44925

Query Match 65.2%; Score 434; DB 5; Length 551;
Best Local Similarity 63.2%; Pred. No. 7.6e-39;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETIELEKTLPEFSRSYWGIRKIGIM 60
||||| : : : : : ||||||| : : : : : ||||||| :
DB 24 WTYHSAENMTYDEASATCQNTHTLVAIONKEIDYLSIDSPSYIWIGIRKYNW 83

OY 61 TWVGTKSLTEBAENWGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117
||||| : : : : : ||||||| : : : : : ||||||| :
DB 84 IWGTHKPLTEGAKMNAPEPNKKQNECDEVEIYIKRPDGTGMWNDERCKKKKHALC 140

RESULT 35
US-09-791-537-121836
Sequence 121836, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
CURRENT APPLICATION NUMBER: US/09/791,537
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 121836
LENGTH: 551
TYPE: PRT
ORGANISM: Oryctolagus cuniculus
US-09-791-537-121836

Query Match 65.2%; Score 434; DB 5; Length 551;
Best Local Similarity 63.2%; Pred. No. 7.6e-39;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETIELEKTLPEFSRSYWGIRKIGIM 60

```
Db      24 WTYHSAENMTYDEASAYCOQNTHTLVAIONKKEIDYLSIDYSPSYWIGIRKYNW 83
      61 TWVGTNKSILTEAEENMGDEPNKKNEKDCVEIYIKRNKDAGKWDACHKRLKALC 117
      84 IWVGTHTKPLTEBAKAKWMBEPNNKKNNEKDCVEIYIKRKPDGTGMWDERCSKKLALC 140

RESULT 36
US-09-791-537-14651
; Sequence 14651, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14651
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-791-537-14651

Query Match
Best Local Similarity 65.2%; Score 434; DB 5; Length 754;
Matches 76; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

Db      27 WTYVYSTRKAVSMYSRIEFCQKHYTDLVAIONKKEIAYLNDVYNSYWMIGIRKINDKW 86
      61 TWVGTNKSILTEAEENMGDEPNKKNEKDCVEIYIKRNKDAGKWDACHKRLKALC 117
      87 TWVGTKRPLETEAEENMGDEPNKKNEKDCVEIYIKSLSPAGKWNDEPCWKRKRLC 143

RESULT 37
US-09-791-537-41672
; Sequence 41672, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41672
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Sus scrofa domestica
US-09-791-537-41672

Query Match
Best Local Similarity 63.8%; Score 425; DB 5; Length 482;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

Db      23 WSYASSTETMTFDDASAYCOQRYTHLVAIONHAIEIYLNSTFNYSASYWMIGIRKINGTW 82
      61 TWVGTNKSILTEAEENMGDEPNKKNEKDCVEIYIKRNKDAGKWDACHKRLKALC 117
      83 TWIGTKKALTEPATWAPGEPPNNKQSNEDCVEIYIKRKDQSGKWNDECSKKLALC 139

RESULT 40
US-09-791-537-93873

Query Match
Best Local Similarity 63.8%; Score 425; DB 5; Length 484;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

Db      23 WSYASSTETMTFDDASAYCOQRYTHLVAIONHAIEIYLNSTFNYSASYWMIGIRKINGTW 82
      61 TWVGTNKSILTEAEENMGDEPNKKNEKDCVEIYIKRNKDAGKWDACHKRLKALC 117
      83 TWIGTKKALTEPATWAPGEPPNNKQSNEDCVEIYIKRKDQSGKWNDECSKKLALC 139

RESULT 38
US-09-791-537-111285
; Sequence 111285, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 111285
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-111285

Query Match
Best Local Similarity 63.2%; Score 425; DB 5; Length 482;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

Db      23 WSYASSTETMTFDDASAYCOQRYTHLVAIONHAIEIYLNSTFNYSASYWMIGIRKINGTW 82
      61 TWVGTNKSILTEAEENMGDEPNKKNEKDCVEIYIKRNKDAGKWDACHKRLKALC 117
      83 TWIGTKKALTEPATWAPGEPPNNKQSNEDCVEIYIKRKDQSGKWNDECSKKLALC 139

RESULT 39
US-09-791-537-50405
; Sequence 50405, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50405
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-50405
```

; Sequence 93873, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 93873
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; US-09-791-537-93873

Query Match 63.1%; Score 420; DB 5; Length 649;
Best Local Similarity 65.0%; Pred. No. 3,1e-37;
Matches 78; Conservative 11; Mismatches 25; Indels 6; Gaps 3;

QY 1 WTYHSEKPMNQARARFCRDNYTDLVATONKAIEIEYLEKTLPSRSYVIGIRK--IG 57
||||| : : ||| : ||||| ||| : ||| : ||||| |
DB 42 WTYHSNKTYSMNYSRAFCQKYTDLVATONKNEIDYLNETIPYNSYVIGIRKQQIN 101
QY 58 GIWTWVGINKSLTEAEENKGDGEPNNKKKEDCVELIYIKRNKDAGKWNDACHKLKALC 117
| ||| ||| ||||| ||||| : ||||| ||||| : |||
DB 102 G--TWVG-NKNLTAEAEENKADNEPNNKRNNDCEVEIYIKSLSAPGKWNDEPCWKRRALC 158

Search completed: September 7, 2002, 10:22:08
Job time: 556 sec

This Page Blank (uspto)

QY 1 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 60
|||||
Db 61 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 120
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKLAALC 117
|||||
Db 121 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKLAALC 177

RESULT 2

US-09-760-443-1328
; Sequence 1328, Application US/09760443
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P212
; CURRENT APPLICATION NUMBER: US/09/760,443
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2164
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1328
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-443-1328

Query Match 100.0%; Score 666; DB 21; Length 341;
Best Local Similarity 100.0%; Pred. No. 9.4e-68;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 60
|||||
Db 61 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 120
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKLAALC 117
|||||
Db 121 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKLAALC 177

RESULT 3

PCT-US01-26675-3
; Sequence 3, Application PC/TUS0126675
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Anastasio, Allison E
; APPLICANT: Bieglecki, Karyn M
; APPLICANT: Klien, Stefanie E
; APPLICANT: Koshy, Beena
; APPLICANT: Kumar, Anant Madan
; TITLE OF INVENTION: HAPLOTYPES OF THE SELL GENE
; FILE REFERENCE: SELL MHI1116-PCT
; CURRENT APPLICATION NUMBER: PCT/US01/26675
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,262
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-26675-3

Query Match 100.0%; Score 666; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 60
|||||
Db 39 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 98
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKLAALC 117
|||||
Db 99 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKLAALC 155

RESULT 4

US-09-119-209-2
; Sequence 2, Application US/09119209
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,209
; FILING DATE: 20-Jul-1998

CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513278
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-119-209-2

Query Match 100.0%; Score 666; DB 15; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 60
|||||
Db 39 WTYHSEKPMNQARRCRDNYTDLVAIONKAEIYEKTLPPSRSYWIGIRKIGIW 98
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKLAALC 117

Db 99 TWVGTNKSILTEAEANWGDSEPNKKKEDCVELIYIKRNKDAGKMWDDACHKRLKALC 155

RESULT 5
US-09-760-475-2123
Sequence 2123, Application US/09760475
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT249
CURRENT APPLICATION NUMBER: US/09/760,475
CURRENT FILING DATE: 2001-01-16
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 4122
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2123
LENGTH: 382
TYPE: PRT
ORGANISM: Homo sapiens
US-09-760-475-2123

Query Match 100.0%; Score 666; DB 21; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSRSYWIIGIRKIGIW 60
Db 49 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSRSYWIIGIRKIGIW 108
Oy 61 TWVGTNKSILTEAEANWGDSEPNKKKEDCVELIYIKRNKDAGKMWDDACHKRLKALC 117
Db 109 TWVGTNKSILTEAEANWGDSEPNKKKEDCVELIYIKRNKDAGKMWDDACHKRLKALC 165

RESULT 6
PCT-US92-03970-2
Sequence 2, Application PC/TUS9203970
GENERAL INFORMATION:
APPLICANT: Dana-Farber Cancer Institute, Inc.
TITLE OF INVENTION: LEUKOCYTE-ASSOCIATED CELL SURFACE
NUMBER OF INVENTIONS: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03970
FILING DATE: 19920513
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Helne, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-152Bq9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: protein
PCT-US92-03970-2

Query Match 100.0%; Score 666; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSRSYWIIGIRKIGIW 60
Db 52 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSRSYWIIGIRKIGIW 111
Oy 61 TWVGTNKSILTEAEANWGDSEPNKKKEDCVELIYIKRNKDAGKMWDDACHKRLKALC 117
Db 112 TWVGTNKSILTEAEANWGDSEPNKKKEDCVELIYIKRNKDAGKMWDDACHKRLKALC 168

RESULT 7
PCT-US94-00909-2
Sequence 2, Application PC/TUS9400909
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS BLOCKING
NUMBER OF INVENTIONS: 11
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (Epo)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00909
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
APPLICATION NUMBER: US 07/962,483
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
APPLICATION DATA:
APPLICATION NUMBER: US 07/737,092
FILING DATE: 29-JUL-1991
APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-00909-2

Query Match 100.0%; Score 666; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSRSYWIIGIRKIGIW 60
Db 52 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSRSYWIIGIRKIGIW 111

OY 61 TWVGNKSLTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDGAKWMDACHKLAALC 117
|||||
Db 112 TWVGNKSLTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDGAKWMDACHKLAALC 168

RESULT 8

US-08-008-459-2
Sequence 2, Application US/08008459
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/008,459
FILING DATE: 25-JAN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/862,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/737,092
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-318XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-008-459-2

Query Match 100.0%; Score 666; DB 4; Length 385;
Best Local Similarity 100.0%; Pred. No. 1,1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WTYHSEKPMNQARRFCRDNYTDVAIONKAEIETLEKTLPSFSRYWIGIRKIGIM 60

Db 52 WTYHSEKPMNQARRFCRDNYTDVAIONKAEIETLEKTLPSFSRYWIGIRKIGIM 111
|||||
OY 61 TWVGNKSLTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDGAKWMDACHKLAALC 117
|||||
Db 112 TWVGNKSLTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDGAKWMDACHKLAALC 168

RESULT 9

US-08-340-539-2
Sequence 2, Application US/08340539
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/862,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/737,092
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-318XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539-2

Query Match 100.0%; Score 666; DB 7; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 60
|||||
DB 52 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 111
|||||
QY 61 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 117
|||||
DB 112 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 168
|||||

RESULT 10

US-08-410-569-2
; Sequence 2, Application US/08410569
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Spertini, Olivier G.
; TITLE OF INVENTION: LEUKOCYTE ADHESION MOLECULE-1 (LAM-1)
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/410,569
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,608
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: US 07/700,773
; FILING DATE: 15-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Helne, Holliday C.
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: DFCG-152EX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; TELEX: 940675
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-410-569-2

Query Match 100.0%; Score 666; DB 8; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 60
|||||
DB 52 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 111
|||||
QY 61 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 117
|||||
DB 112 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 168
|||||

RESULT 11
US-60-207-315-428
; Sequence 428, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-428

Query Match 100.0%; Score 666; DB 26; Length 1078;
Best Local Similarity 100.0%; Pred. No. 4.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 60
|||||
DB 621 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 680
|||||
QY 61 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 117
|||||
DB 681 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 737
|||||

RESULT 12
US-60-212-659-523
; Sequence 523, Application US/60212659
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1000674
; CURRENT APPLICATION NUMBER: US/60/212,659
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 879
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-212-659-523

Query Match 100.0%; Score 666; DB 26; Length 1078;
Best Local Similarity 100.0%; Pred. No. 4.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 60
|||||
DB 621 WTYHSEKPMQORARRCRDNYTDLVAIONKAEIYELEKTLPSRSYYWIGIRKIGTW 680
|||||
QY 61 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 117
|||||
DB 681 TWGYNKSLTEAEENWGDEPNKKKEDCVEIYIKRNKDAGKWDACCHKIKALC 737
|||||

RESULT 13
US-60-230-435-1751
; Sequence 1751, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: USES THEREOF

```

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1751
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1751
```

```

Query Match          100.0%; Score 666; DB 26; Length 1078;
Best Local Similarity 100.0%; Pred. No. 4.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 60
    |||||||
DB 621 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 680
    |||||||
QY 61 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 117
    |||||||
DB 681 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 737
    |||||||
```

```

RESULT 14
US-60-160-189-8687
; Sequence 8687, Application US/60160189
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; FILE REFERENCE: CL000112
; CURRENT APPLICATION NUMBER: US/60/160,189
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 10162
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8687
; LENGTH: 119
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-189-8687
```

```

Query Match          98.2%; Score 654; DB 26; Length 119;
Best Local Similarity 98.3%; Pred. No. 6e-67;
Matches 115; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 60
    |||||||
DB 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 60
    |||||||
QY 61 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 117
    |||||||
DB 61 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 117
    |||||||
```

```

RESULT 15
US-60-169-867-5823
; Sequence 5823, Application US/60169867
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: CL000160
```

```

; CURRENT APPLICATION NUMBER: US/60/169,867
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 8230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5823
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-867-5823
```

```

Query Match          98.2%; Score 654; DB 26; Length 119;
Best Local Similarity 98.3%; Pred. No. 6e-67;
Matches 115; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 60
    |||||||
DB 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 60
    |||||||
QY 61 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 117
    |||||||
DB 61 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 117
    |||||||
```

```

RESULT 16
US-60-160-203-5003
; Sequence 5003, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5003
; LENGTH: 119
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-203-5003
```

```

Query Match          95.0%; Score 633; DB 26; Length 119;
Best Local Similarity 94.9%; Pred. No. 1.6e-64;
Matches 111; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 60
    |||||||
DB 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYYWIGIRKIGIW 60
    |||||||
QY 61 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 117
    |||||||
DB 61 TWVGINKSLTEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALC 117
    |||||||
```

```

RESULT 17
US-60-169-840-6716
; Sequence 6716, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
```

```
; FILE REFERENCE: CLO00164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6716
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-840-6716

Query Match          95.0%; Score 633; DB 26; Length 119;
Best Local Similarity 94.9%; Pred. No. 1.6e-64;
Matches 111; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWIIGIRKIGIW 60
    |||||||
DB 1 WYHYSEKPMNQRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWIIGIRKIGIW 60
    |||||||

QY 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDAGKWNDACHKLKAALC 117
    |||||||
DB 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDAGKWNDACHKLKAALC 117
    |||||||

RESULT 18
US-09-119-209-4
; Sequence 4, Application US/09119209
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,209
; FILING DATE: 20-Jul-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513278
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-119-209-4

Query Match          84.8%; Score 565; DB 15; Length 372;
Best Local Similarity 82.9%; Pred. No. 4.8e-56;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWIIGIRKIGIW 60
    |||||||
DB 39 WYHYSEKPMNQRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWIIGIRKIGIW 98
    |||||||

QY 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDAGKWNDACHKLKAALC 117
    |||||||
DB 99 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDAGKWNDACHKLKAALC 155
    |||||||

RESULT 19
US-60-160-189-10011
; Sequence 10011, Application US/60160189
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00112
; CURRENT APPLICATION NUMBER: US/60/160,189
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 10162
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10011
; LENGTH: 116
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-189-10011

Query Match          80.9%; Score 539; DB 26; Length 116;
Best Local Similarity 96.0%; Pred. No. 1.1e-53;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWIIGIRKIGIW 60
    |||||||
DB 11 WYHYSEKPMNQRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWIIGIRKIGIW 70
    |||||||

QY 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDAGKWNDACHKLKAALC 117
    |||||||
DB 71 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDAGKWNDACHKLKAALC 155
    |||||||

RESULT 20
US-60-160-203-6200
; Sequence 6200, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 6200
; LENGTH: 116
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-203-6200
```

```
Query Match
Best Local Similarity 80.9%; Score 539; DB 26; Length 116;
Best Local Similarity 96.0%; Pred. No. 1.1e-53;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||
DB 11 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 70
OY 61 TWGTTNKSLTEEAENMGDEPNKKKEDCVEIYIKRK 99
    |||
DB 71 TWGTTNKSLTEEAENMGDEPNKKKEDCVEIYIKRKQ 109
```

```
RESULT 21
US-60-169-840-9326
; Sequence 9326, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9326
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-840-9326
```

```
Query Match
Best Local Similarity 80.9%; Score 539; DB 26; Length 116;
Best Local Similarity 96.0%; Pred. No. 1.1e-53;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||
DB 11 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 70
OY 61 TWGTTNKSLTEEAENMGDEPNKKKEDCVEIYIKRK 99
    |||
DB 71 TWGTTNKSLTEEAENMGDEPNKKKEDCVEIYIKRKQ 109
```

```
RESULT 22
US-60-169-867-7998
; Sequence 7998, Application US/60169867
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C1000160
; CURRENT APPLICATION NUMBER: US/60/169,867
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 8230
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7998
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-867-7998
```

```
Query Match
Best Local Similarity 80.9%; Score 539; DB 26; Length 116;
Best Local Similarity 96.0%; Pred. No. 1.1e-53;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||
DB 11 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 70
OY 61 TWGTTNKSLTEEAENMGDEPNKKKEDCVEIYIKRK 99
    |||
DB 71 TWGTTNKSLTEEAENMGDEPNKKKEDCVEIYIKRKQ 109
```

```
RESULT 23
US-60-207-315-467
; Sequence 467, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 467
; LENGTH: 700
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(700)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-207-315-467
```

```
Query Match
Best Local Similarity 66.8%; Score 445; DB 26; Length 700;
Best Local Similarity 65.8%; Pred. No. 7.5e-42;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;
```

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||
DB 34 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 93
OY 61 TWGTTNKSLTEEAENMGDEPNKKKEDCVEIYIKRKADGKKNDDACHLKAALC 117
    |||
DB 94 TWGTTNKSALTEEAENMGDEPNKKKEDCVEIYIKRSPADGKNWDEHCKLKAALC 150
```

```
RESULT 24
PCT-US94-09395-4
; Sequence 4, Application PC/TUS9409395
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: Expression Control Sequences of the P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
```



```

; FILE REFERENCE: CL000456
; CURRENT APPLICATION NUMBER: US/60/196,718
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4236
; LENGTH: 129
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-718-4236

```

Query Match

```

Best Local Similarity 66.1%; Score 440.5; DB 26; Length 129;
Matches 77; Conservative 13; Mismatches 26; Indels 1; Gaps 1;

```

```

QY 1 WTYHSEKPMNMOARARFCRDNYTDLVAIONKAEIELEYLEKTLPEFSRSYWGIRKIGTW 60
    ||||| | : : : : : ||||| | : || : ||||| |
Db 6 WTYHSTKAYSWE-SRKYCQNRDYLVAIONKNEIDLNVLPYSSYWGIRKNNKTW 65
QY 61 TWVGTNKSLEAEANWGDGEPNNKKNEDEVEIYIKRNKDAGKWNDACHKLKAALC 117
    ||||| | : ||||| | : ||||| | : ||||| | : ||||| | : ||||| |
Db 66 TWVGTKALITNEAEANWADNEPNKKRNNEDEVEIYIKESV-SGKWNDEHCLKKHALC 121

```

RESULT 28

```

US-60-196-718-4237
; Sequence 4237, Application US/60196718
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000456
; CURRENT APPLICATION NUMBER: US/60/196,718
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4237
; LENGTH: 133
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-718-4237

```

Query Match

```

Best Local Similarity 66.1%; Score 440.5; DB 26; Length 133;
Matches 77; Conservative 13; Mismatches 26; Indels 1; Gaps 1;

```

```

QY 1 WTYHSEKPMNMOARARFCRDNYTDLVAIONKAEIELEYLEKTLPEFSRSYWGIRKIGTW 60
    ||||| | : : : : : ||||| | : || : ||||| |
Db 17 WTYHSTKAYSWE-SRKYCQNRDYLVAIONKNEIDLNVLPYSSYWGIRKNNKTW 76
QY 61 TWVGTNKSLEAEANWGDGEPNNKKNEDEVEIYIKRNKDAGKWNDACHKLKAALC 117
    ||||| | : ||||| | : ||||| | : ||||| | : ||||| | : ||||| |
Db 77 TWVGTKALITNEAEANWADNEPNKKRNNEDEVEIYIKESV-SGKWNDEHCLKKHALC 132

```

RESULT 29

```

US-60-195-053-1909
; Sequence 1909, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1909
; LENGTH: 128

```

```

; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-1909

```

Query Match

```

Best Local Similarity 66.0%; Score 439.5; DB 26; Length 128;
Matches 77; Conservative 12; Mismatches 27; Indels 1; Gaps 1;

```

```

QY 1 WTYHSEKPMNMOARARFCRDNYTDLVAIONKAEIELEYLEKTLPEFSRSYWGIRKIGTW 60
    ||||| | : : : : : ||||| | : || : ||||| |
Db 11 WTYHSTKAYSWE-SRKYCQNRDYLVAIONKNEIDLNVLPYSSYWGIRKNNKTW 69
QY 61 TWVGTNKSLEAEANWGDGEPNNKKNEDEVEIYIKRNKDAGKWNDACHKLKAALC 117
    ||||| | : ||||| | : ||||| | : ||||| | : ||||| | : ||||| |
Db 70 TWVGTKALITNEAEANWADNEPNKKRNNEDEVEIYIKSPAPGKWNDEHCLKKHALC 126

```

RESULT 30

```

US-60-195-053-1908
; Sequence 1908, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1908
; LENGTH: 129
; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-1908

```

Query Match

```

Best Local Similarity 66.0%; Score 439.5; DB 26; Length 129;
Matches 77; Conservative 12; Mismatches 27; Indels 1; Gaps 1;

```

```

QY 1 WTYHSEKPMNMOARARFCRDNYTDLVAIONKAEIELEYLEKTLPEFSRSYWGIRKIGTW 60
    ||||| | : : : : : ||||| | : || : ||||| |
Db 6 WTYHSTKAYSWE-SRKYCQNRDYLVAIONKNEIDLNVLPYSSYWGIRKNNKTW 64
QY 61 TWVGTNKSLEAEANWGDGEPNNKKNEDEVEIYIKRNKDAGKWNDACHKLKAALC 117
    ||||| | : ||||| | : ||||| | : ||||| | : ||||| | : ||||| |
Db 65 TWVGTKALITNEAEANWADNEPNKKRNNEDEVEIYIKSPAPGKWNDEHCLKKHALC 121

```

RESULT 31

```

US-60-196-718-4238
; Sequence 4238, Application US/60196718
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000456
; CURRENT APPLICATION NUMBER: US/60/196,718
; CURRENT FILING DATE: 2000-04-13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4238
; LENGTH: 138
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-718-4238

```

```

Query Match
Best Local Similarity 66.0%; Score 439.5; DB 26; Length 138;
Matches 77; Conservative 12; Mismatches 27; Indels 1; Gaps 1;

```



```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-657-753-2

Query Match
Best Local Similarity 62.3%; Score 415; DB 10; Length 610;
Matches 71; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNMQRARFCRDNYTDLVAIONKAEIETLEKTLFPSRSYWIIGIRKIGIW 60
DB 22 WSYNTSTAMTYDEASAYCQQRTHLVAIQNKEEIEYLNSTLISYSPSYWIGIRKVNWW 81
QY 61 TWVGTKSLTEBAENWGDEPNKRNKEDCVETIYIKRNKDAGKWNDDACHKLAALC 117
DB 82 VVVGTOKPLTEBAKNWAPGEPNNRQKDEDCVETIYIKREKDVGMWNDERCKSKLALC 138

RESULT 36
US-08-770-435-3
Sequence 3, Application US/08770435
GENERAL INFORMATION:
APPLICANT: Gimbrone, Jr., Michael A.
APPLICANT: Bevilacqua, Michael P.
TITLE OF INVENTION: Antihodies Specific For E-selectin And The Uses
TITLE OF INVENTION: thereof
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 New York Ave., NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/770,435
FILING DATE: herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/365,470
FILING DATE: 29-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,510
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/850,802
FILING DATE: 13-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Matkovic, Karen R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 0627.1350004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-770-435-3
```

```
Query Match
Best Local Similarity 62.3%; Score 415; DB 11; Length 610;
Matches 71; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNMQRARFCRDNYTDLVAIONKAEIETLEKTLFPSRSYWIIGIRKIGIW 60
DB 22 WSYNTSTAMTYDEASAYCQQRTHLVAIQNKEEIEYLNSTLISYSPSYWIGIRKVNWW 81
QY 61 TWVGTKSLTEBAENWGDEPNKRNKEDCVETIYIKRNKDAGKWNDDACHKLAALC 117
DB 82 VVVGTOKPLTEBAKNWAPGEPNNRQKDEDCVETIYIKREKDVGMWNDERCKSKLALC 138

RESULT 37
US-09-266-091-2
Sequence 2, Application US/09266091
GENERAL INFORMATION:
APPLICANT: Klimuk, Sandra K
APPLICANT: Semple, Sean C
APPLICANT: Scherrer, Peter
APPLICANT: Hope, Michael J.
APPLICANT: Bennett, C. Frank
TITLE OF INVENTION: ENHANCED EFFICACY OF LIPOSOMAL ANTISENSE THERAPY
FILE REFERENCE: ISPH-0342
CURRENT APPLICATION NUMBER: US/09/266,091
CURRENT FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 08/657,753
PRIOR FILING DATE: 1996-05-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 610
TYPE: PRT
ORGANISM: Homo sapiens
US-09-266-091-2

Query Match
Best Local Similarity 62.3%; Score 415; DB 16; Length 610;
Matches 71; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNMQRARFCRDNYTDLVAIONKAEIETLEKTLFPSRSYWIIGIRKIGIW 60
DB 22 WSYNTSTAMTYDEASAYCQQRTHLVAIQNKEEIEYLNSTLISYSPSYWIGIRKVNWW 81
QY 61 TWVGTKSLTEBAENWGDEPNKRNKEDCVETIYIKRNKDAGKWNDDACHKLAALC 117
DB 82 VVVGTOKPLTEBAKNWAPGEPNNRQKDEDCVETIYIKREKDVGMWNDERCKSKLALC 138

RESULT 38
US-09-266-091A-2
Sequence 2, Application US/09266091A
GENERAL INFORMATION:
APPLICANT: Klimuk, Sandra K.
APPLICANT: Semple, Sean C.
APPLICANT: Scherrer, Peter
APPLICANT: Hope, Michael J.
TITLE OF INVENTION: Enhanced Efficacy of Liposomal Antisense
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:10:07 : Search time 28.63 Seconds
(without alignments)
99.818 Million cell updates/sec

Title: US-09-119-209-2_COPY_39_155

Perfect score: 666

Sequence: 1 WTYHYSEKPMNMQRARFCR.....NKDAGKWNDDACHKRAALC 117

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	117	1	US-08-274-661B-38 Sequence 38, Appl
2	666	100.0	372	2	US-08-513-278-2 Sequence 2, Appl
3	666	100.0	372	6	5514582-2 Patent No. 5514582
4	666	100.0	385	1	US-08-340-539A-2 Sequence 2, Appl
5	666	100.0	385	2	US-08-461-592B-2 Sequence 2, Appl
6	575	86.3	119	1	US-08-340-539A-14 Sequence 14, Appl
7	565	84.8	117	6	5514582-7 Patent No. 5514582
8	565	84.8	372	2	US-08-513-278-4 Sequence 4, Appl
9	565	84.8	372	6	5514582-4 Patent No. 5514582
10	560	84.1	119	1	US-08-340-539A-12 Sequence 12, Appl
11	557	83.6	119	1	US-08-340-539A-13 Sequence 13, Appl
12	455	67.1	119	1	US-08-340-539A-18 Sequence 18, Appl
13	447	67.1	830	5	PCF-US91-05059-2 Sequence 2, Appl
14	447	67.1	830	6	5378464-2 Patent No. 5378464
15	445	66.8	117	1	US-08-274-661B-39 Sequence 39, Appl
16	445	66.8	119	1	US-08-340-539A-16 Sequence 16, Appl
17	445	66.8	119	1	US-08-340-539A-17 Sequence 17, Appl
18	445	66.8	830	1	US-08-110-158-4 Sequence 4, Appl
19	444	66.7	120	1	US-08-340-539A-19 Sequence 19, Appl
20	434	65.2	120	1	US-08-274-661B-37 Sequence 37, Appl
21	425	63.8	484	2	US-08-252-493C-9 Sequence 9, Appl
22	425	63.8	484	3	US-09-276-197-9 Sequence 9, Appl
23	415	62.3	119	1	US-08-340-539A-15 Sequence 15, Appl
24	415	62.3	120	1	US-08-274-661B-36 Sequence 36, Appl
25	415	62.3	610	1	US-08-365-470-3 Sequence 3, Appl
26	415	62.3	610	3	US-09-209-668-19 Sequence 19, Appl
27	415	62.3	610	4	US-09-009-490A-89 Sequence 89, Appl

28	415	62.3	610	6	5217870-2 Patent No. 5217870
29	237	35.6	67	3	US-08-840-062-8 Sequence 8, Appl
30	164	24.6	38	2	US-08-513-278-5 Sequence 5, Appl
31	164	24.6	38	6	5514582-5 Patent No. 5514582
32	161	24.2	31	2	US-08-140-137A-44 Sequence 44, Appl
33	160.5	24.1	273	4	US-09-111-470-10 Sequence 10, Appl
34	160.5	24.1	292	2	US-08-688-342-4 Sequence 4, Appl
35	160.5	24.1	292	2	US-09-113-788-4 Sequence 4, Appl
36	160.5	24.1	316	4	US-09-111-470-4 Sequence 4, Appl
37	154.5	23.2	110	6	5514582-12 Patent No. 5514582
38	154.5	23.2	174	1	US-07-641-971B-1 Sequence 1, Appl
39	154.5	23.2	174	1	US-07-781-248A-1 Sequence 1, Appl
40	154.5	23.2	320	1	US-08-365-103B-10 Sequence 10, Appl
41	154.5	23.2	321	1	US-08-365-103B-8 Sequence 8, Appl
42	150.5	22.6	287	4	US-09-111-470-6 Sequence 6, Appl
43	149.5	22.4	114	6	5514582-8 Patent No. 5514582
44	149	22.4	111	6	5514582-11 Patent No. 5514582
45	148	22.2	1479	3	US-08-840-062-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-08-274-661B-38

Sequence 38, Application US/08274661B

Patent No. 5593882

GENERAL INFORMATION:

APPLICANT: Erbe, David V.

APPLICANT: Lasky, Laurence A.

APPLICANT: Presta, Leonard G.

TITLE OF INVENTION: Selectin Variants

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESS: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/274,661B

FILING DATE: 13-Jul-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/956701

FILING DATE: 10/01/1992

ATTORNEY/AGENT INFORMATION:

NAME: Dregler, Ginger R.

REGISTRATION NUMBER: 33,055

REFERENCE/DOCKET NUMBER: 761P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-3216

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids

TOPOLOGY: Linear

US-08-274-661B-38

Query Match 100.0%, Score 666, DB 1, Length 117;

Best Local Similarity 100.0%, Pred. No. 1.8e-66;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 1 WTYHYSEKPMNMQRARFCRDNTDVAIONKAEIYLEKTLPEFSRSYIWIGIRKIGIWM 60

Db 1 WTHYSEKPMWQRRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 60
QY 61 TWGTNKSLEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWNDACHKLKAALC 117
Db 61 TWGTNKSLEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWNDACHKLKAALC 117

RESULT 2
US-08-513-278-2
; Sequence 2, Application US/08513278
; Patent No. 5840844
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patln (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,278
; FILING DATE: 10-AUG-1993
; CLASSIFICATION: 5530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 06-MAY-1993
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 565D1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-513-278-2

Query Match 100.0%; Score 666; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 7.5e-66;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTHYSEKPMWQRRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 60
Db 39 WTHYSEKPMWQRRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 98
QY 61 TWGTNKSLEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWNDACHKLKAALC 117
Db 99 TWGTNKSLEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWNDACHKLKAALC 155

RESULT 3
5514582-2
; Patent No. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
; IMMNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO:2:
; LENGTH: 372
5514582-2

Query Match 100.0%; Score 666; DB 6; Length 372;
Best Local Similarity 100.0%; Pred. No. 7.5e-66;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTHYSEKPMWQRRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 60
Db 39 WTHYSEKPMWQRRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 98
QY 61 TWGTNKSLEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWNDACHKLKAALC 117
Db 99 TWGTNKSLEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWNDACHKLKAALC 155

RESULT 4
US-08-340-539A-2
; Sequence 2, Application US/08340539A
; Patent No. 5808025
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,539A
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,459
; FILING DATE: 25-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gunnison, Jane
; REGISTRATION NUMBER: 38,479
; REFERENCE/DOCKET NUMBER: CG-104 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-2

Query Match 100.0%; Score 666; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 7.8e-66;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAETLEKTLPEFSRSYWIIGIRKIGTW 60
|||||
Db 52 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAETLEKTLPEFSRSYWIIGIRKIGTW 111
Qy 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWDACHKRLKALC 117
|||||
Db 112 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWDACHKRLKALC 168

RESULT 5
US-08-461-592B-2
Sequence 2, Application US/08461592B
Patent No. 5834425

GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,592B
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/340,539
FILING DATE: 16-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: CG-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-592B-2

Query Match 100.0%; Score 666; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 7.8e-66;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAETLEKTLPEFSRSYWIIGIRKIGTW 60
|||||
Db 52 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAETLEKTLPEFSRSYWIIGIRKIGTW 111
Qy 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWDACHKRLKALC 117
|||||
Db 112 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWDACHKRLKALC 168

RESULT 6
US-08-340-539A-14
Sequence 14, Application US/08340539A
Patent No. 5808025

GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9090
TELEFAX: 212-596-9000
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-14

Query Match 86.3%; Score 575; DB 1; Length 119;
Best Local Similarity 82.1%; Pred. No. 2.1e-56;
Matches 96; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

Qy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAETLEKTLPEFSRSYWIIGIRKIGTW 60
|||||
Db 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAETLEKTLPEFSRSYWIIGIRKIGTW 60
Qy 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWDACHKRLKALC 117
|||||
Db 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWDACHKRLKALC 117

RESULT 7
5514582-7
Patent No. 5514582

APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 7
LENGTH: 117
5514582-7

Query Match 84.8%; Score 565; DB 6; Length 117;
Best Local Similarity 82.9%; Pred. No. 2.7e-55;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAEIETLEKTLPSRSYYWIGIRKIGIW 60
DB 1 WTYHSEKPMNMENARKFCOKNTYDLVAIONKREIETLENTLPKSPYYWIGIRKIGKW 60
QY 61 TWGINKSLTEAEENWGGEPPNNKKNEKDCVEIYIKRKNKDGKWNDDACHKRAALC 117
DB 61 TWGINKTLTKEAENWGAGEPPNNKKSKEDCVEIYIKRERDSGKWNDDACHKRAALC 117

RESULT 8
US-08-513-278-4
Sequence 4, Application US/08513278
Patent No. 5840844
GENERAL INFORMATION:
APPLICANT: LASKY, LAURENCE A.
APPLICANT: STACHELL, SCOTT E.
APPLICANT: ROSEN, STEVEN D.
APPLICANT: SINGER, MARK S.
APPLICANT: YEDNOCK, TED A.
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,278
FILING DATE: 10-AUG-1995
CLASSIFICATION: 5530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/059027
FILING DATE: 06-MAY-1993
APPLICATION NUMBER: 07/786149
FILING DATE: 31-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/315015
FILING DATE: 23-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055

REFERENCE/DOCKET NUMBER: 565DICI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-513-278-4

Query Match 84.8%; Score 565; DB 2; Length 372;
Best Local Similarity 82.9%; Pred. No. 1.1e-54;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAEIETLEKTLPSRSYYWIGIRKIGIW 60
DB 39 WTYHSEKPMNMENARKFCOKNTYDLVAIONKREIETLENTLPKSPYYWIGIRKIGKW 98
QY 61 TWGINKSLTEAEENWGGEPPNNKKNEKDCVEIYIKRKNKDGKWNDDACHKRAALC 117
DB 99 TWGINKTLTKEAENWGAGEPPNNKKSKEDCVEIYIKRERDSGKWNDDACHKRAALC 155

RESULT 9
5514582-4
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 4
LENGTH: 372
5514582-4

Query Match 84.8%; Score 565; DB 6; Length 372;
Best Local Similarity 82.9%; Pred. No. 1.1e-54;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAEIETLEKTLPSRSYYWIGIRKIGIW 60
DB 39 WTYHSEKPMNMENARKFCOKNTYDLVAIONKREIETLENTLPKSPYYWIGIRKIGKW 98
QY 61 TWGINKSLTEAEENWGGEPPNNKKNEKDCVEIYIKRKNKDGKWNDDACHKRAALC 117
DB 99 TWGINKTLTKEAENWGAGEPPNNKKSKEDCVEIYIKRERDSGKWNDDACHKRAALC 155
RESULT 10
US-08-340-539A-12
Sequence 12 Application US/08340539A
Patent No. 5608025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTING AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-12

Query Match      84.1%; Score 560; DB 1; Length 119;
Best Local Similarity 82.1%; Pred. No. 9.7e-55;
Matches 96; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRARFCRDNTDLVAIONKAEIYLEKTLFSPSYWIGIRKIGTW 60
    |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 1 WYHYSEKPMNQRARFCRDNTDLVAIONKAEIYLEKTLFSPSYWIGIRKIGTW 60
QY 61 TWVGINKSLTEAEKNGDEPNKKKNEKCEVEIYIKRNNDAGKWNDDACHKRAALC 117
    |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 61 TWVGINKSLTEAEKNGDEPNKKKNEKCEVEIYIKRNNDAGKWNDDACHKRAALC 117
US-08-340-539A-13
RESULT 11
Sequence 13, Application US/08340539A
Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-13
```

```

FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-13

Query Match      83.6%; Score 557; DB 1; Length 119;
Best Local Similarity 81.2%; Pred. No. 2.1e-54;
Matches 95; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRARFCRDNTDLVAIONKAEIYLEKTLFSPSYWIGIRKIGTW 60
    |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 1 WYHYSEKPMNQRARFCRDNTDLVAIONKAEIYLEKTLFSPSYWIGIRKIGTW 60
QY 61 TWVGINKSLTEAEKNGDEPNKKKNEKCEVEIYIKRNNDAGKWNDDACHKRAALC 117
    |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 61 TWVGINKSLTEAEKNGDEPNKKKNEKCEVEIYIKRNNDAGKWNDDACHKRAALC 117
US-08-340-539A-18
RESULT 12
Sequence 18, Application US/08340539A
Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-18
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
US-08-340-539A-13
RESULT 11
Sequence 13, Application US/08340539A
Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-13
```

```

FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-18
RESULT 12
Sequence 18, Application US/08340539A
Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-18
```


NAME/KEY: Disulfide-bond
LOCATION: 572..616
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 585..616
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 602..629
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 642..686
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 655..668
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 672..699
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 704..748
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 717..730
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 734..761
FEATURE:
NAME/KEY: Binding-site
LOCATION: 54
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 98
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 180
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 212
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 219
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 411
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 460
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 518
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 665
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site

LOCATION: 716
OTHER INFORMATION: /note= "Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
Query Match 67.1%; Score 447; DB 5; Length 830;
Best Local Similarity 65.8%; Pred. No. 3.2e-41;
Matches 77: Conservative 11; Mismatches 29; Indels 0; Gaps 0;
OY 1 WYHYSEKPMQRRRCRDNYTDLVAIQNAEIELEKTLPEFSRSTYWGIRKIGIW 60
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHYSTRAYSWNISRKYCQNRVYTDVAIQNNEIDYINKVLPYSSYWGIRKNNWTW 101
OY 61 TWGVTNKSLEAEENMGDEPNKKKEDCVETIYKRNKDGKWNDDACHKLAALC 117
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGTRKALTNEAEWADNEPNKKRNEDCVETIYKSPAPGKWNDECKLKKMAALC 158
RESULT 14
5378464-2
PATENT NO. 5378464
APPLICANT: MCEVER, RODGER P.
TITLE OF INVENTION: MODULATION OF INFLAMMATORY RESPONSES
BY ADMINISTRATION OF GMP-140 OR ANTIBODY TO GMP-140
NUMBER OF SEQUENCES: 32
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/320,408
FILING DATE: 08-MAR-1989
SEQ ID NO.: 2
LENGTH: 830
5378464-2
Query Match 67.1%; Score 447; DB 6; Length 830;
Best Local Similarity 65.8%; Pred. No. 3.2e-41;
Matches 77: Conservative 11; Mismatches 29; Indels 0; Gaps 0;
OY 1 WYHYSEKPMQRRRCRDNYTDLVAIQNAEIELEKTLPEFSRSTYWGIRKIGIW 60
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHYSTRAYSWNISRKYCQNRVYTDVAIQNNEIDYINKVLPYSSYWGIRKNNWTW 101
OY 61 TWGVTNKSLEAEENMGDEPNKKKEDCVETIYKRNKDGKWNDDACHKLAALC 117
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGTRKALTNEAEWADNEPNKKRNEDCVETIYKSPAPGKWNDECKLKKMAALC 158
RESULT 15
US-08-274-661B-39
Sequence 39, Application US/08274661B
Patent No. 5593882
GENERAL INFORMATION:
APPLICANT: Erbe, David V.
APPLICANT: Lasky, Laurence A.
APPLICANT: Presta, Leonard G.
TITLE OF INVENTION: Selection Variants
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/274,661B
FILING DATE: 13-Jul-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

```
QY      1 WUYHSEKEMNNQRARFRCDNYDYLVAIQNKAEIYELEKTLPFSRSYWIIGRIJGIGIW 60
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 WUYHYSTKAYSNNISRKYCOMRNEDYLVAIONKNNEIDYLNKFLPYSZYWYGIRKNKWT 60
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
        61 TWVGTKRSLTEEAENMGSGEPNKKKNCDCVEYITIKRKDAGCKKNDACIKLKAALC 117
       ||||| |:| | ||||| | ||||| :| ||||| :| | ||||| :| | |||||
```

Db 61 TWVGTAKLTNEAENWADNEPNNKRNEDCVELIYKSPSPGKWNDEHCKRRKHALC 117

RESULT 18

US-08-110-158-4
; Sequence 4, Application US/08110158
; Patent No. 5605821
; GENERAL INFORMATION:

APPLICANT: McEver, Rodger P.

APPLICANT: Pan, Junliang

TITLE OF INVENTION: Expression Control Sequences of the

TITLE OF INVENTION: P-Selectin Gene

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 1100 Peachtree Street, Suite 2800

CITY: Atlanta

STATE: GA

COUNTRY: USA

ZIP: 30309-4530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/110,158

FILING DATE: 19930820

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/320,408

FILING DATE: 08-MAR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404)-815-6508

TELEFAX: (404)-815-6555

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 830 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-110-158-4

Query Match 66.8%; Score 445; DB 1; Length 830;
Best Local Similarity 65.8%; Pred. No. 5.4e-41;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRARFRCDNYTDLVAIONKAEIYLEKTLPFSSSYWIGIRKIGTW 60
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHSTAYSNISRKVCQNRYTDLVAIONKNEIDYLNKVPYSSYWIGIRKNNKW 101
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 61 TWVGTAKLTNEAENWADNEPNNKRNEDCVELIYKSPSPGKWNDEHCKRRKHALC 117
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWVGTAKLTNEAENWADNEPNNKRNEDCVELIYKSPSPGKWNDEHCKRRKHALC 158

RESULT 19
US-08-340-539A-19
; Sequence 19, Application US/08340539A
; Patent No. 5808025
; GENERAL INFORMATION:

APPLICANT: Tedder, Thomas F.

APPLICANT: Kansas, Geoffrey S.

TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS

TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,539A

FILING DATE: 16-NOV-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/008,459

FILING DATE: 25-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Gunnison, Jane

REGISTRATION NUMBER: 38,479

REFERENCE/DOCKET NUMBER: CG-104 CON

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 119 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-340-539A-19

Query Match 66.7%; Score 444; DB 1; Length 119;
Best Local Similarity 65.8%; Pred. No. 6.4e-42;
Matches 77; Conservative 10; Mismatches 30; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRARFRCDNYTDLVAIONKAEIYLEKTLPFSSSYWIGIRKIGTW 60
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 1 WTYHSEKPMNQRARFRCDNYTDLVAIONKNEIDYLNKVPYSSYWIGIRKNNKW 60
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 61 TWVGTAKLTNEAENWADNEPNNKRNEDCVELIYKSPSPGKWNDEHCKRRKHALC 117
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 61 TWVGTAKLTNEAENWADNEPNNKRNEDCVELIYKSPSPGKWNDEHCKRRKHALC 117

RESULT 20
US-08-274-661B-37
; Sequence 37, Application US/08274661B
; Patent No. 5593882
; GENERAL INFORMATION:

APPLICANT: Erbe, David V.

APPLICANT: Lasky, Laurence A.

APPLICANT: Presta, Leonard G.

TITLE OF INVENTION: Selectin Variants

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/274,661B

FILING DATE: 13-Jul-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956701
FILING DATE: 10/01/1992
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: Amino acid
TOPOLOGY: Linear
US-08-274-661B-37

Query Match 65.2%; Score 434; DB 1; Length 120;
Best Local Similarity 63.2%; Pred. No. 8.3e-41;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTHYSEKPMWQRRRCRNDTDLVAIONKAEIYEKTLPPSRSYTWIGIRKIGTW 60
DB 1 WTHYSEKPMWQRRRCRNDTDLVAIONKAEIYEKTLPPSRSYTWIGIRKIGTW 60
OY 61 TWGTSKSLTEAEENWGDGEPNNKKNKEDCEYIYIKRNKDGKWNDDACHKKAALC 117
DB 61 IWGTHKFLTEGAKNMAGPEPNKNNQNNEDCEYIYIKRKDGKWNDDCHSKKLAALC 117

RESULT 21
US-08-252-493C-9
Sequence 9, Application US/08252493C
Patent No. 5891645
GENERAL INFORMATION:
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Evans, Mark J.
APPLICANT: Matlis, Louis A.
TITLE OF INVENTION: PORCINE E-SELECTIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park, Box 15
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 Kb storage
COMPUTER: PC compatible
OPERATING SYSTEM: DOS 6.2
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,493C
FILING DATE: June 1, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid

STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: predicted amino acid sequence of
DESCRIPTION: Porcine E-selectin
US-08-252-493C-9

Query Match 63.8%; Score 425; DB 2; Length 484;
Best Local Similarity 63.2%; Pred. No. 4.5e-39;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTHYSEKPMWQRRRCRNDTDLVAIONKAEIYEKTLPPSRSYTWIGIRKIGTW 60
DB 23 WSYASTETWTFEDASAYCOQRYTHLVAIONHAEIYELNSTFVNSASYTWIGIRKIGTW 82
OY 61 TWGTSKSLTEAEENWGDGEPNNKKNKEDCEYIYIKRNKDGKWNDDACHKKAALC 117
DB 83 TWIGTKALTEPATNMAGPEPNKKNKEDCEYIYIKRKDGKWNDDCHSKKLAALC 139

RESULT 22
US-09-276-197-9
Sequence 9, Application US/09276197
Patent No. 6040428
GENERAL INFORMATION:

APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Evans, Mark J.
APPLICANT: Matlis, Louis A.
TITLE OF INVENTION: PORCINE E-SELECTIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park, Box 15
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 Kb storage
COMPUTER: PC compatible
OPERATING SYSTEM: DOS 6.2
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,197
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/252,493
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: predicted amino acid sequence of
DESCRIPTION: Porcine E-selectin
US-09-276-197-9

Query Match 63.8%; Score 425; DB 3; Length 484;
Best Local Similarity 63.2%; Pred. No. 4.5e-39;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

APPLICANT: CARON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 5:
LENGTH: 38
5514582-5

Query Match 24.6%; Score 164; DB 6; Length 38;
Best Local Similarity 78.4%; Pred. No. 1.4e-11;
Matches 29; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 TYHYSEKPMNMOARARFCRDNTDVAIONKAEIETLPS 38
DB 2 TYHYSEKPMNMOARARFCRDNTDVAIONKAEIETLPS 38

RESULT 32
US-08-140-137A-44
Sequence 44, Application US/08140137A
Patent No. 5817617
GENERAL INFORMATION:
APPLICANT: TUOMANEN, ELAINE
TITLE OF INVENTION: ANALOGS OF ENDOTHELIAL LEUKOCYTE
ADHESION MOLECULE (ELAM)
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/140,137A
FILING DATE: 27-MAY-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-096
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: hLHRc peptide sequence
US-08-140-137A-44

Query Match 24.2%; Score 161; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.4e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 ARRFCDNTDVAIONKAEIETLPS 45
DB 1 ARRFCDNTDVAIONKAEIETLPS 31

RESULT 33
US-09-111-470-10
Sequence 10, Application US/09111470
Patent No. 6277959
GENERAL INFORMATION:
APPLICANT: Valladeau, Jenny
APPLICANT: Ravel, Odile
APPLICANT: Bates, Elizabeth E.M.
APPLICANT: Ford, John
APPLICANT: Saeland, Sem
APPLICANT: Lebecque, Serge J.E.
TITLE OF INVENTION: Mammalian Membrane Protein Genes;
TITLE OF INVENTION: Related Reagents
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,470
FILING DATE: 08-JUL-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/053,080
FILING DATE: 09-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0695
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-111-470-10

Query Match 24.1%; Score 160.5; DB 4; Length 273;
Best Local Similarity 29.6%; Pred. No. 3.9e-10;
Matches 34; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

OY 3 YHYSEKPMNMOARARFCRDNTDVAIONKAEIETLPSFSYWIIGRIIGITW 62
DB 150 YWFSHSGMSMAEAIEYQOLKNAHLVINSRQONFVOKYL--GSAIYWGSLSPDEGAKW 207
OY 63 V-GTRKSLTEEAENWGDGEPNNKRNK-----EDCVETIYIRNKDAGKAMDAGCHK 111
DB 208 VDGTD--YATGFQWKKPGQPDMDQGHGIGGEDCAHFH-----PDGRMNDVQGR 255

RESULT 34


```

US-08-688-342-4
; Sequence 4, Application US/08688342
; Patent No. 5871964
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goll, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,342
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CTP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1235724
US-08-688-342-4

Query Match          24.1%; Score 160.5; DB 2; Length 292;
Best Local Similarity 29.6%; Pred. NO. 4.2e-10;
Matches 34; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

QY 3 YHYSEKPNMWORARRFCDNDYTDLVAIONKAEIEYLEKTLPEFSRSYVWIGIRKIGITW 62
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 169 YWFSHSGSMWADEAKRYCQCKNAHLLVINSREQNFVQKYL--GSAVYTWGGLSDPGAKWK 226
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 63 V-GTAKSLTELEENKGDGPNNKKK-----EDCYETITKRNKDKAGKNDACRK 111
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 227 VDGTD--YATGFQNNKPGQPPDDMOGHGIGGEDCAHFH-----PDGRWMDVDCOR 274
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 35
US-09-113-788-4
; Sequence 4, Application US/09113788
; Patent No. 5969104
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goll, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto

```

```

: STATE: CA
: COUNTRY: US
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/113,788
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/688,342
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 292 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: IMMEDIATE SOURCE:
: LIBRARY: Genbank
: CLONE: 1235724
:
: US-09-113-788-4
:
: Query Match 24.1%; Score 160.5; DB 2; Length 292;
: Best Local Similarity 29.6%; Pred. No. 4.2e-10;
: Matches 34; Conservative 22; Mismatches 44; Indels 15; Gaps 5;
:
: QY 3 YHSEKPMWOFARFCNDNYTLVAIONKAEIEYLEKTLPSRSYVYIGIRKIGIWTW 62
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
: Db 169 YMFHSIGSMKAEAEYCCQKNAHLVINSREQNFVQKYL--GSAVTWNGLSDPGAMKW 226
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
: QY 63 V-GTKSLTEAEKNGDEPNKKKK-----EDCVEIYIKRNKDGAKWDDACHK 111
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
: Db 227 VDGTD--YATGFQNKPGQRPDDMQHGIGGEGDCAHFH-----PDGRWDDVQCR 274
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
:
: RESULT 36
: US-09-111-470-4
: Sequence 4, Application US/09111470
: Patent No. 6277959
: GENERAL INFORMATION:
: APPLICANT: Valladeau, Jenny
: APPLICANT: Ravel, Odile
: APPLICANT: Bates, Elizabeth E.M.
: APPLICANT: Ford, John
: APPLICANT: Saeland, Sem
: APPLICANT: Lebecque, Serge J.E.
: TITLE OF INVENTION: Mammalian Membrane Protein Genes:
: TITLE OF INVENTION: Related Reagents
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DNAX Research Institute
: STREET: 901 California Avenue
: CITY: Palo Alto
: STATE: California
: COUNTRY: USA
: ZIP: 94304-1104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-111-470-4

```

```

Query Match          24.1%; Score 160.5; DB 4; Length 316;
Best Local Similarity 29.6%; Pred. No. 4.6e-10;
Matches 34; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

```

```

QY 3 YHSEKPMNORARRCRDNYTDVAIONKAEIEYLEKTLPSRSYWGIRKIGITW 62
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 193 YFSGHSMGMAEAEKCYCOLKNAHLVINSREONVQKYL--GSAYTWGSLDPGCAWK 250
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 63 V-GTKSKLSEAEENNGDEPNKKKK-----EDCVEIYIKRNKDGAKWDDACRK 111
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 251 VDGTD--YATGFQNMKPPQPDMDGHLGGEDCAHFH-----PDGRNNDVDCQR 298
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 37
5514582-12
; PATENT NO. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
; IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO: 12:
; LENGTH: 110
5514582-12

```

```

Query Match          23.2%; Score 154.5; DB 6; Length 110;
Best Local Similarity 32.2%; Pred. No. 5.8e-10;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

```

```

QY 3 YHSEKPMNORARRCRDNYTDVAIONKAEIEYLEKTLPSRSYWGIRKIG--GIW 60
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 3 YFGKGTGKQVHARACDDMEGLVSIHSPEDFLTKHSHSGS--WIGLRNLDLKGEP 60
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 61 TWVGTKSLSEAEENNGDEPNKKKKEDCVEIYIKRNKDGAKWDDAC--HKLKALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 61 IWDGSH--VDISNMAPGEPTSRSGEDCVAM-----RSGRWMDAFCDRKLGAMVC 110
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 38
US-07-641-971B-1
; Sequence 1, Application US/07/641971B
; Patent No. 5236706
; GENERAL INFORMATION:
; APPLICANT: Debire, Patrice
; APPLICANT: Mossalayi, Mohammed D
; TITLE OF INVENTION: A PHARMACEUTICAL PREPARATION FOR THE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Irving M. Fishman, CIBA-GEIGY Corporation
; STREET: 556 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/641,971B
; FILING DATE: 19910116
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 90016254
; FILING DATE: 24-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fishman, Irving M
; REGISTRATION NUMBER: 30258
; REFERENCE/DOCKET NUMBER: 4-17921/+DEB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-277-4832
; TELEFAX: 908-277-4306
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; CELL TYPE: Human B. Cells
; CELL LINE: CHO cells transformed with pCAL8-BF-ND
; US-07-641-971B-1

```

```

Query Match          23.2%; Score 154.5; DB 1; Length 174;
Best Local Similarity 32.2%; Pred. No. 1e-09;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

```

```

QY 3 YHSEKPMNORARRCRDNYTDVAIONKAEIEYLEKTLPSRSYWGIRKIG--GIW 60
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 28 YFGKGTGKQVHARACDDMEGLVSIHSPEDFLTKHSHSGS--WIGLRNLDLKGEP 85
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 61 TWVGTKSLSEAEENNGDEPNKKKKEDCVEIYIKRNKDGAKWDDAC--HKLKALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 86 IWDGSH--VDISNMAPGEPTSRSGEDCVAM-----RSGRWMDAFCDRKLGAMVC 135
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 39
US-07-781-248A-1
; Sequence 1, Application US/07/781248A
; Patent No. 5246699
; GENERAL INFORMATION:
; APPLICANT: Debire, Patrice
; APPLICANT: Mossalayi, Mohammed D
; TITLE OF INVENTION: MATURATION OF HEMATOPOIETIC CELLS
; NUMBER OF SEQUENCES: 6

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Irving M. Fishman, CIBA-GEIGY Corporation
STREET: 556 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/81,248A
FILING DATE: 19911230
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 90103565
FILING DATE: 09-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Ikeler, Barbara J.
REGISTRATION NUMBER: 36,170
REFERENCE/DOCKET NUMBER: 4-18065/A/DEB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-277-3368
TELEFAX: 908-277-4306
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
CELL TYPE: Human B. Cells
CELL LINE: CHO cells transformed with pCAL8-BF-ND
US-07-781-248A-1

Query Match          23.2%; Score 154.5; DB 1; Length 174;
Best Local Similarity 32.2%; Pred. No. 1e-09;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

QY 3 YHSEKPMNWARPCRDNYDVAIONKAEIEYLEKTLPPSRSYWIGIRKIG--GITW 60
   1::: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 28 YFEGKTQWVHARYACDMEGOLVSHSPEDQFLTKHASHTGS--WIGLRNLDLKGEE 85
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

QY 61 TWVGTKSLTEAEKMGDEPNKKKKEDCEVEIYIKRNKADGKWNDDAC-HKTKAALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 86 IWDGSH---VDYSNMAPGEPTSRSGEDCVMM-----RSGRWNDAPCDRLGAWVC 135
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 40
US-08-365-103B-10
Sequence 10 Application US/08365103B
Patent No. 576943
GENERAL INFORMATION:
APPLICANT: Lynch, Richard G
APPLICANT: Nunez, Raphael D.
APPLICANT: Yodot, Jungi
TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thombe, Voorhees & Sease
STREET: 801 Grand Ave. Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,103B
FILING DATE: 28-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: Ulf N5-24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 288-3667
TELEFAX: (515) 288-1338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-365-103B-10

Query Match          23.2%; Score 154.5; DB 1; Length 320;
Best Local Similarity 32.2%; Pred. No. 2.2e-09;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

QY 3 YHSEKPMNWARPCRDNYDVAIONKAEIEYLEKTLPPSRSYWIGIRKIG--GITW 60
   1::: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 174 YFEGKTQWVHARYACDMEGOLVSHSPEDQFLTKHASHTGS--WIGLRNLDLKGEE 231
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

QY 61 TWVGTKSLTEAEKMGDEPNKKKKEDCEVEIYIKRNKADGKWNDDAC-HKTKAALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 232 IWDGSH---VDYSNMAPGEPTSRSGEDCVMM-----RSGRWNDAPCDRLGAWVC 281
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

Search completed: September 7, 2002, 10:14:42
Job time: 275 sec

This Page Blank (uspto)


```
FT Modified-site 324..326
FT /label= N-linked_glycosylation
FT /note= "putative"
XX
XX
XX WO9201049-A.
XX
XX 23-JAN-1992.
XX
XX 15-JUL-1990; 90WO-US04986.
XX
XX 13-JUL-1990; 90US-0553759.
XX
XX (GEHO-) GEN HOSPITAL CORP.
XX
XX Seed B, Arufio A, Amiot M;
XX
XX WPI: 1992-056864/07.
XX N-PSDB; AAQ22500.
XX
XX New CD53 cell surface antigen and DNA encoding it - for
XX immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
XX Example 14; Page 106; 160pp; English.
XX
XX Two cDNA clones encoding Leu8 determinants were isolated from a
XX human T cell library using the rapid immunoselection cloning method
XX (see e.g. AAQ21164 for description of method). This protein sequence
XX was deduced from the shorter insert. The weakly hydrophobic
XX C-terminal domain is characteristic of surface proteins that are
XX attached to the cell membrane by covalent linkage to a
XX phosphatidylinositol-substituted glycan.
XX See AAQ21184 for the larger insert and AAR20815 for the major form of
XX the Leu8 antigen that it encodes.
XX
XX Sequence 363 AA;
XX
XX
XX Query Match 100.0%; Score 666; DB 13; Length 363;
XX Best Local Similarity 100.0%; Pred. No. 2,2e-63;
XX Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 WTYHYSEKPMNMOARARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWGIRKIGIW 60
XX |||||||
XX Db 52 wtyhysekpmnwqarrirfrtdyldvalgnkaeylektlpfsrsywygirkigiw 111
XX |||||||
XX QY 61 TWVGSTNKSLTTEEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDGAKWMDACHIKKALC 117
XX |||||||
XX Db 112 twvgtnkslteeeenwgdgeppnnkknkedcveiyikrnkdagkwnddachiklkaalc 168
XX |||||||
XX
XX RESULT 2
XX AAR91443
XX ID AAR91443 standard; Protein; 363 AA.
XX
XX AAR91443;
XX
XX 31-OCT-1996 (first entry)
XX
XX Human Leu8 antigen.
XX
XX Cell surface antigen; cloning; immunoselection; immunotherapy;
XX therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;
XX antiinflammatory.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 113..115
XX /label= Glycosylation_site
XX Modified-site 156..158
XX /label= Glycosylation_site
XX Modified-site 229..231
XX /label= Glycosylation_site
XX
```

```
FT Modified-site 269..271
FT /label= Glycosylation_site
FT 285..287
FT Modified-site /label= Glycosylation_site
FT 299..301
FT Modified-site /label= Glycosylation_site
FT 324..326
FT /label= Glycosylation_site
XX
XX US506126-A.
XX
XX 09-APR-1996.
XX
XX 25-FEB-1988; 88US-0160416.
XX
XX 01-DEC-1992; 92US-0983647.
XX 25-FEB-1988; 88US-0160416.
XX 13-JUL-1989; 89US-0379076.
XX 13-JUL-1990; 90US-0553759.
XX 18-OCT-1993; 93US-0139273.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Arufio A, Seed B;
XX
XX WPI: 1996-200279/20.
XX N-PSDB; AAT14723.
XX
XX Cloning of cDNA encoding cell surface antigen - useful for isolation
XX of diagnostic and therapeutic proteins
XX
XX Example 14; Column 71-74; 79pp; English.
XX
XX
XX 2 cDNA clones encoding Leu8 determinants (AAR91442 and AAR91443) were
XX isolated from a human T-cell library using a novel immunoselection
XX cloning method. The longer insert (AAT14723) contained 2,350
XX residues, while the shorter lacked 436 internal residues. A major
XX transcript of 2.4 kb was present in peripheral blood mononuclear
XX cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,
XX and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukemia
XX T-cell lines. Leu8 antigens can be obtd. for diagnostic and
XX therapeutic use. The presence or absence of Leu8 on CD4+ T-cells
XX identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.
XX Soluble forms of Leu8 can act as antiinflammatory agents by reducing
XX lymphocyte migration.
XX
XX Sequence 363 AA;
XX
XX
XX Query Match 100.0%; Score 666; DB 17; Length 363;
XX Best Local Similarity 100.0%; Pred. No. 2,2e-63;
XX Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 WTYHYSEKPMNMOARARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWGIRKIGIW 60
XX |||||||
XX Db 52 wtyhysekpmnwqarrirfrtdyldvalgnkaeylektlpfsrsywygirkigiw 111
XX |||||||
XX QY 61 TWVGSTNKSLTTEEAEENMGDGEPPNNKKNKEDCVEIYIKRNKDGAKWMDACHIKKALC 117
XX |||||||
XX Db 112 twvgtnkslteeeenwgdgeppnnkknkedcveiyikrnkdagkwnddachiklkaalc 168
XX |||||||
XX
XX RESULT 3
XX AAY96184
XX ID AAY96184 standard; Protein; 363 AA.
XX
XX AAY96184;
XX
XX 19-DEC-2000 (first entry)
XX
XX Human T-cell specific Leu8 antigen.
XX
XX Leu8; cell surface antigen; human; immunoselection; panning;
XX
```

KW immunodiagnosis; diagnosis; immunotherapy; gene therapy;
KW immune disorder; infection; aschma; immune-complex disease;
KW amyloidosis; multiple sclerosis; inflammation; antinflammatory.

OS Homo sapiens.

PN US6111093-A.

XX 29-AUG-2000.

PD 28-OCT-1998; 98US-0181612.

PF 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 23-MAR-1990; 90US-0498809.

XX 13-JUL-1990; 90US-0553759.

XX (GEHO) GEN HOSPITAL CORP.

PI Stamenkovic I, Seed B;

XX WPI; 2000-586382/55.

DR N-PSDB; AAA50632.

XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
PT useful for immunodiagnosis and immunotherapy of immune-mediated
PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
PT diseases

XX Example 14; Column 69-72; 75pp: English.

XX The present sequence is that of a T-cell specific Leu8 antigen, as
CC predicted from 1 of 2 clones (see AAA50632) isolated from a human
CC T-cell library by the method of the invention. A longer Leu8
CC antigen is given in AA96138. The method, designed to isolate cell
CC surface antigen (CSA) cDNAs, is based upon transient expression of
CC a CSA in eukaryotic cells and physical selection of cells expressing
CC the antigen by adhesion to (panning on) an antibody-coated substrate
CC such as a culture dish. CSA nucleic acids isolated by the method of
CC the invention, and the proteins they encode, are useful for
CC immunodiagnosis and immunotherapeutic applications, including the
CC diagnosis and treatment of immune-mediated infections, diseases, and
CC disorders in animals, including humans. These disorders include
CC asthma, immune-complex disease, amyloidosis, parasitic diseases or
CC multiple sclerosis. The ability to interfere with the binding of
CC Leu8-T-cells to antigen presenting cells, or the ability to cause
CC such binding to occur on surfaces other than lymphocyte cells, can
CC be useful in diagnostics and therapy. The level of activated Leu8-
CC T-cells relative to resting Leu8+ cells could serve as a measure of
CC immune response to a particular antigen. Modification of the
CC specificity of the extracellular domain of Leu8, which mediates
CC adhesion to specific endothelial cells of lymph nodes, could serve
CC to regulate the homing potential of resting T cells. Soluble forms
CC of Leu8 could act as antiinflammatory agents by reducing lymphocyte
CC migration.

XX Sequence 363 AA:

Query Match 100.0%; Score 666; DB 21: Length 363;
Best Local Similarity 100.0%; Pred. No. 2.2e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQRRRCRDNYTDVAIONKAETLEKTLPFNSRYWIGIRKIGIM 60

DB 52 WYHSEKPMNQRRRCRDNYTDVAIONKAETLEKTLPFNSRYWIGIRKIGIM 111

OY 61 TWVGINKSLTEAEKMGDEPNKKKKCEVETIYIKRNKDAGKWDNDACHKIKAALC 117

DB 112 TWVGINKSLTEAEKMGDEPNKKKKCEVETIYIKRNKDAGKWDNDACHKIKAALC 168

RESULT 4

AA98127
ID AAR98127 standard; Protein; 370 AA.

XX AAR98127;

DT 01-NOV-1996 (first entry)

XX Human lymphocyte cell surface glycoprotein (HULHR) variant.

XX Immunoglobulin; transmembrane receptor; adhesion; targeting;

KW immunodiagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

XX metastasis.

OS Homo sapiens.

XX Key

XX Region

FT Domain

FT Domain

FT Binding-site

FT Binding-site

FT Domain

FT Domain

FT Domain

XX US5514582-A.

XX 07-MAY-1996.

XX 23-FEB-1989; 89US-0315015.

XX 22-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

XX (GEHTH) GENENTECH INC.

PA Capon DJ, Lasky LA;

XX WPI; 1996-238773/24.

XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand

PT binding site of a receptor fused to Ig constant region - useful for

PT diagnosis and treatment e.g. of inflammation

XX Disclosure; Page 19; 41pp: English.

XX A hybrid immunoglobulin chain comprising the ligand binding site of

CC a single transmembrane receptor without an active transmembrane

CC region; fused at its C-terminus with the N-terminus of an

CC immunoglobulin constant region. The receptor is not a member of the

CC immunoglobulin super family, nor a multiple subunit polypeptide

CC encoded by discrete genes. The hybrid immunoglobulin chain combines

CC the adhesion/targeting of a ligand binding partner (LBP) with the

CC effector functions of immunoglobulin and can bind to and/or activate

CC more than one ligand. It can be used diagnostically for the in

CC vitro assay of LBP and their targets; or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.

CC Typical applications are as antiviral, neuromodulating and

CC immunomodulating agents, or as modulators of cell adhesion (e.g. in

CC treating graft rejection; inflammation; metastasis of lymphoma etc.)

CC The immunoglobulin component increases plasma half life and

CC facilitates purification while deletion of the transmembrane region

FT /Label= Cytoplasmic domain.
XX US5514582-A.
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH) GENENTECH INC.
XX Capon DJ, Lasky LA;
XX WPI: 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
PS Disclosure: Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant is an Ile296 deletion mutant.
XX
XX
SQ Sequence 370 AA;

Query Match 100.0%; Score 666; DB 17; Length 370;
Best Local Similarity 100.0%; Pred. No. 2,3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVHYSEKPMQMMORARRCBDNYTDLVAIONKAEIEYEKTPFSRSYYWIGIRKIGIW 60
Db 38 WTVHYSEKPMQMMORARRCBDNYTDLVAIONKAEIEYEKTPFSRSYYWIGIRKIGIW 97
QY 61 TWVGINKSLTEAEANWGDGEPNNKKNKEDCYEYITKRNKDGAKMNDCHKLKALC 117
Db 98 TWVGINKSLTEAEANWGDGEPNNKKNKEDCYEYITKRNKDGAKMNDCHKLKALC 154

RESULT 7
AAR98114
ID AAR98114 standard; Protein: 371 AA.
XX
AC AAR98114;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX

KW Immunoglobulin: transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FT Region 1..37
FT /Label= Signal region.
FT Domain 38..154
FT /Label= Lectin domain.
FT Domain 159..192
FT /Label= EGF domain.
FT Binding-site 196..257
FT /Label= Complement binding repeat 1.
FT Binding-site 258..316
FT /Label= Complement binding repeat 2.
FT Domain 332..354
FT /Label= Transmembrane domain.
FT Domain 355..371
FT /Label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX WPI: 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
PS Disclosure: Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an His168Gln substitution.
XX
XX
SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 2,3e-63;

[illegible]

CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets, or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases metastasis half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an Ile174Leu substitution.
XX

Sequence 371 AA;

Query Match	100.0%;	Score 666;	DB 17;	Length 371;
Best Local Similarity	100.0%;	Pred. No. 2.3e-63;		
Matches 117;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 1 WYHSEKPMNQRRRCNDNYDVAIONKAIEIYELEKTPEFSRSYYWIGIRIKGIW 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 38 wtyhysekpmmqrrarrferdndydlvalnqkaeiylektlpfsrissywigirikiqiwiw 977

Qy 61 TWVGTSKSLTEEAENWGDGEPNNKKNKEDCEYIYIKRNKDAGKWNDDACHKLEKALC 117
|||||
Db 98 twvgtskslteeaenwgdgepnknknkedceyiyikrnkdagkwnddachlkaalc 154

RESULT	9
AAR98116	
ID	AAR98116 standard; Protein; 371 AA

AC	AAR98116;
XX	
DT	01-NOV-1996 (first entry)

DE Human lymphocyte cell surface glycoprotein (HLIR) variant
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targetting

diagnosis; therapy; drug delivery; antiviral; neuromodulator; immunomodulator; cell adhesion; graft rejection; inflammation; metastasis.

OS Homo sapiens.

FH	key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
FT
/label= Signal region.
```

/label= Lectin domain.

/label= EGF domain.

```

/label= Complement binding repeat 1.

```

/label= Complement binding repeat 2.

/label= Transmembrane domain.

/label= Cytoplasmic domain.

PN US5514582-A.

PD 07-MAY-1996.

PF 23-FEB-1989; 89US-0315015.
XX

PR 22-NOV-1989; 89US-0440625.

PR 22-NOV-1989;
PR 23-FEB-1989;

PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX
XX WPI: 1996-238773/24.
XX
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX
XX Disclosure: Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an Asn181Gln substitution.
XX
XX
SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRRCRDNYTDLVAIQNAELEYLEKTPFSRSYWIIRKIGIM 60
Db 38 WYHYSEKPMNQRRRCRDNYTDLVAIQNAELEYLEKTPFSRSYWIIRKIGIM 97
Oy 61 TWVGTKNSLTFEAEKMGDEPNKKKNEKDCVEIYIKRNKDAGKWNDDACHKIKALC 117
Db 98 TWVGTKNSLTFEAEKMGDEPNKKKNEKDCVEIYIKRNKDAGKWNDDACHKIKALC 154

RESULT 10
AAR98117
ID AAR98117 standard; Protein: 371 AA.
XX
XX AAR98117;
XX
XX 01-NOV-1996 (first entry)
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Immunoglobulin: transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..37
FT Region
FT /label= Signal region.

FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371
FT /label= Cytoplasmic domain.
PN US514582-A.
PD 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX
XX WPI: 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX
XX Disclosure: Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Thr211Ser substitution.
XX
XX
SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRRCRDNYTDLVAIQNAELEYLEKTPFSRSYWIIRKIGIM 60
Db 38 WYHYSEKPMNQRRRCRDNYTDLVAIQNAELEYLEKTPFSRSYWIIRKIGIM 97
Oy 61 TWVGTKNSLTFEAEKMGDEPNKKKNEKDCVEIYIKRNKDAGKWNDDACHKIKALC 117
Db 98 TWVGTKNSLTFEAEKMGDEPNKKKNEKDCVEIYIKRNKDAGKWNDDACHKIKALC 154

```

RESULT 11
AA98118
ID AAR98118 standard; Protein; 371 AA.
XX
AC AAR98118;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371
FT /label= Cytoplasmic domain.
XX
PN US514582-A.
XX
PD 07-MAY-1996.
XX
PE 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family; nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and

```

```

CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Phe214Leu substitution.
XX
SQ Sequence 371 AA;
XX
Query Match 100.0%; Score 666; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WTYHSEKPMNQRARRECDNTDVAIONKAEIYLEKTPPSSTYWTGIRKIGGTM 60
DQ 38 WYHYSEKPMNQRARRECDNTDVAIONKAEIYLEKTPPSSTYWTGIRKIGGTM 60
DQ 61 TWGTSNKLTEAEANMGDGPNNKKNKEDCEVEIYIKRKNKDGAKWMDACHTKKAALC 117
DQ 98 twvgtnkslteeenwgdgpnknkknkedceveiyikrknkdagkwmdacthikaalc 154
XX
RESULT 12
AA98119
ID AAR98119 standard; Protein; 371 AA.
XX
AC AAR98119;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371
FT /label= Cytoplasmic domain.
XX
PN US514582-A.
XX
PD 07-MAY-1996.
XX
PE 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1996-238773/24.
XX

```

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure: Page 19, 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets, or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Ser226Thr substitution.
XX
SQ Sequence 371 AA:

Query Match 100.0%; Score 666; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WYHYSEKPMNQRRRCRDNYDVLAIQNKAELEYLEKTLPSRSYVWIGIRKIGTW 60
Db 38 WYHYSEKPMNQRRRCRDNYDVLAIQNKAELEYLEKTLPSRSYVWIGIRKIGTW 97
Qy 61 TWVGYNKSLTEFEANWGDGPENKKNKEDCVELYIKRNKDGAKMDDACHIKALC 117
Db 98 TWVGYNKSLTEFEANWGDGPENKKNKEDCVELYIKRNKDGAKMDDACHIKALC 154

RESULT 13
AAR98120
ID AAR98120 standard; Protein: 371 AA.
AC AAR98120;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Region 1..37
FT Domain /label= Signal region.
FT Domain 38..154
FT Domain /label= Lectin domain.
FT Domain 159..192
FT Domain /label= EGF domain.
FT Binding-site 196..257
FT Binding-site /label= Complement binding repeat 1.
FT Binding-site 258..316
FT Binding-site /label= Complement binding repeat 2.
FT Domain 332..354
FT Domain /label= Transmembrane domain.

FT Domain 355..371
FT /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
PD 07-MAY-1996.
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GENE) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI, 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure: Page 19, 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets, or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Phe244Met substitution.
XX
SQ Sequence 371 AA:

Query Match 100.0%; Score 666; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WYHYSEKPMNQRRRCRDNYDVLAIQNKAELEYLEKTLPSRSYVWIGIRKIGTW 60
Db 38 WYHYSEKPMNQRRRCRDNYDVLAIQNKAELEYLEKTLPSRSYVWIGIRKIGTW 97
Qy 61 TWVGYNKSLTEFEANWGDGPENKKNKEDCVELYIKRNKDGAKMDDACHIKALC 117
Db 98 TWVGYNKSLTEFEANWGDGPENKKNKEDCVELYIKRNKDGAKMDDACHIKALC 154

RESULT 14
AAR98121
ID AAR98121 standard; Protein: 371 AA.
AC AAR98121;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

```

XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371
FT /label= Cytoplasmic domain.
FT Domain
XX
PN US514582-A.
XX
PD 07-MAY-1996.
XX
PE 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI: 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery. Improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Thr282Ser substitution.
XX
SQ Sequence 371 AA;

```

```

Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WTVHYSEKPMNORARFRCDNYTDLVAIONKAETLEKTLFPFSRYWIGIRKIGTAW 60
Db 38 Wlyhysekpmnwqrrarfrcrdnydlvaigkneylektpfsrywigirkigtw 97
OY 61 TWVGTNKSITFEARNWDDGEPNNKKKEDCEVEIYIKRNKDAGKWDACHKKAALC 117
Db 98 twvgtlnkslteeenwgdgepnkkkedcveiyikrnkdagkwndachkklkaalc 154

RESULT 15
AAR98122
ID AAR98122 standard; Protein; 371 AA.
XX
AC AAR98122;
XX
DE 01-NOV-1996 (first entry)
XX
XX Human lymphocyte cell surface glycoprotein (HuLHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371
FT /label= Cytoplasmic domain.
XX
PN US514582-A.
XX
PD 07-MAY-1996.
XX
PE 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI: 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the

```

CC Immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC the adhesion/targeting of a ligand binding partner (LBP) with the
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver
 CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery, improves aqueous solubility and removes
 CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-989135. This variant contains an Ile288Val substitution.
 CC
 XX
 SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNQARARFCRDNYTDLVAIONKAEIYELEKTLPFSSRYWIGIRKIGTW 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 38 WYHYSEKPMNQARARFCRDNYTDLVAIONKAEIYELEKTLPFSSRYWIGIRKIGTW 97
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 61 TWVGTKSLTEEAENWGDEPNKKKEDCEVEIYIRKNKDGAKWMDACHKIKALC 117
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 98 TWVGTKSLTEEAENWGDEPNKKKEDCEVEIYIRKNKDGAKWMDACHKIKALC 154
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 16
 AAR98123
 ID AAR98123 standard; Protein; 371 AA.
 AC AAR98123;
 XX
 DT 01-NOV-1996 (first entry)
 XX
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
 XX
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 KM
 OS Homo sapiens.
 XX
 XX

Key Location/Qualifiers
 FT Region 1..37
 FT Label= Signal region.
 FT Domain 38..154
 FT Label= Lectin domain.
 FT Domain 159..192
 FT Label= EGF domain.
 FT Binding-site 196..257
 FT Label= Complement binding repeat 1.
 FT Binding-site 258..316
 FT Label= Complement binding repeat 2.
 FT Domain 332..354
 FT Label= Transmembrane domain.
 FT Domain 355..371
 FT Label= Cytoplasmic domain.

XX
 XX US5514582-A.
 XX PN 07-MAY-1996.
 XX PD 23-FEB-1989; 89US-0315015.
 XX PF 22-NOV-1989; 89US-0440625.
 XX PR

PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.
 XX
 XX (GENENTECH INC.
 PA Capon DJ, Lasky LA;
 PI WPL; 1996-238773/24.
 DR
 XX
 XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
 PR binding site of a receptor fused to Ig constant region - useful for
 PT diagnosis and treatment e.g. of inflammation
 XX
 PS Disclosure: Page 19, 41pp; English.
 CC
 CC A hybrid immunoglobulin chain comprising the ligand binding site of
 CC a single transmembrane receptor without an active transmembrane
 CC region; fused at its C-terminus with the N-terminus of an
 CC immunoglobulin constant region. The receptor is not a member of the
 CC immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC the adhesion/targeting of a ligand binding partner (LBP) with the
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver
 CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery, improves aqueous solubility and removes
 CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-989135. This variant contains a Lys298-Lys299; Arg-Arg
 CC substitution.
 CC
 XX
 SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNQARARFCRDNYTDLVAIONKAEIYELEKTLPFSSRYWIGIRKIGTW 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 38 WYHYSEKPMNQARARFCRDNYTDLVAIONKAEIYELEKTLPFSSRYWIGIRKIGTW 97
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 61 TWVGTKSLTEEAENWGDEPNKKKEDCEVEIYIRKNKDGAKWMDACHKIKALC 117
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 98 TWVGTKSLTEEAENWGDEPNKKKEDCEVEIYIRKNKDGAKWMDACHKIKALC 154
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 17
 AAR98124
 ID AAR98124 standard; Protein; 371 AA.
 AC AAR98124;
 XX
 DT 01-NOV-1996 (first entry)
 XX
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
 XX
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 KM
 OS Homo sapiens.
 XX
 XX

Key Location/Qualifiers

CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery, improves aqueous solubility and removes
 CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-R98135. This variant is a Asn271 deletion mutant.
 CC
 XX
 SQ Sequence 371 AA;
 Query Match 100.0%; Score 666; DB 17; Length 371;
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WTYHYSEKPMNQARARFCRDNYTDLVATONKAETLEYEKTLPFSRSYWGIRKIGTW 60
 DB 39 WCYHYSEKPMNQARARFCRDNYTDLVATONKAETLEYEKTLPFSRSYWGIRKIGTW 98
 QY 61 TWVGTNKSITBEAENMGDEPNKKRNKEDCEVEIYIKRNKDAGKWDACHKLAALC 117
 DB 99 TWVGTNKSITBEAENMGDEPNKKRNKEDCEVEIYIKRNKDAGKWDACHKLAALC 155
 RESULT 19
 AAR37960
 ID AAR37960 standard; Protein; 372 AA.
 XX
 AC AAR37960;
 DT 08-OCT-1993 (first entry)
 XX
 DE Human Lymphocyte Homing Receptor.
 XX
 KM HuLHR; lymphocyte binding inhibition; lymphoma metastasis;
 KM transplant rejection; inflammation.
 KM
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..38
 FT Protein /label= signal_sequence
 FT 39..372
 FT Domain /note= "Trp39 is probable N-terminus of mature LHR"
 FT 39..155
 FT Modified-site /label= Lectin_domain
 FT 60..62
 FT Modified-site /note= "potential N-glycosylation site"
 FT 104..106
 FT Modified-site /note= "potential N-glycosylation site"
 FT 160..193
 FT Domain /label= EGF_domain
 FT 177..179
 FT Modified-site /note= "potential N-glycosylation site"
 FT 197..258
 FT Region /label= Complement_Binding_Repeat_1
 FT 216..218
 FT Modified-site /note= "potential N-glycosylation site"
 FT 232..234
 FT Modified-site /note= "potential N-glycosylation site"
 FT 246..248
 FT Modified-site /note= "potential N-glycosylation site"
 FT 259..317
 FT Region /label= Complement_Binding_Repeat_2
 FT 271..273
 FT Modified-site /note= "potential N-glycosylation site"
 FT 311..313
 FT Modified-site /note= "potential N-glycosylation site"
 FT 333..335
 FT Domain /label= Transmembrane_Domain
 FT 356..372
 FT Domain /label= Cytoplasmic_Domain

XX
 PN US5216131-A.
 XX
 PD 01-JUN-1993.
 XX
 PF 23-FEB-1989; 89US-0315015.
 XX
 PR 23-FEB-1989; 89US-0315015.
 PR 31-OCT-1991; 91US-0786149.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Lasky LA, Rosen SD, Singer MS, Stachel SE, Yednock TA;
 XX WPI: 1993-188588/23.
 DR N-PSDB: AAQ43154.
 XX
 PT Human and murine lymphocyte homing receptors to treat graft
 PT rejection and inflammation - comprise carbohydrate binding,
 PT epidermal growth factor and complement binding domains
 PS Claim 1: Fig 1 and Fig 3; 32pp; English.
 XX
 CC A human peripheral blood lymphocyte cDNA library in lambda gt10 was
 CC screened with a 2.2kb EcoRI insert of the murine Mel14 antigen clone
 CC (i.e. a murine LHR sequence). The largest EcoRI insert (2.2kb) was
 CC isolated and sequenced. The ORF codes for 372 amino acids with a mol.
 CC wt. of approximately 42,200. Comparison of the HuLHR amino acid
 CC sequence with the murine LHR sequence (AAR37961) showed a high degree
 CC of amino acid conservation in each of the LHR domains, e.g. 96% in
 CC the transmembrane domain and 83% in the carbohydrate binding domain.
 CC The LHRs could be used to compete with the normal binding of
 CC lymphocytes to lymphoid tissue to treat inflammation or graft
 CC rejection. They could also be used to control lymphoma metastasis
 CC and to treat conditions involving lymphocyte accumulation.
 XX
 SQ Sequence 372 AA;
 Query Match 100.0%; Score 666; DB 14; Length 372;
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WTYHYSEKPMNQARARFCRDNYTDLVATONKAETLEYEKTLPFSRSYWGIRKIGTW 60
 DB 39 WCYHYSEKPMNQARARFCRDNYTDLVATONKAETLEYEKTLPFSRSYWGIRKIGTW 98
 QY 61 TWVGTNKSITBEAENMGDEPNKKRNKEDCEVEIYIKRNKDAGKWDACHKLAALC 117
 DB 99 TWVGTNKSITBEAENMGDEPNKKRNKEDCEVEIYIKRNKDAGKWDACHKLAALC 155
 RESULT 20
 AAR38908
 ID AAR38908 standard; Protein; 372 AA.
 XX
 AC AAR38908;
 DT 11-JAN-1994 (first entry)
 XX
 DE HuLHR.
 XX
 KM Human; murine; lymphocyte; cell surface glycoprotein; homing receptor;
 KM LHR; endothelium; lymphoid tissue; signal; domain; complement binding;
 KM carbohydrate binding; epidermal growth factor-like; egf; intracellular;
 KM transmembrane binding; cytoplasmic; ligand binding partner protein;
 KM TMD; LBPP.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 20..32
 FT /note= "Signal peptide"

```

FT Protein 39..372
FT /note= "Mature protein"
FT Domain 39..155
FT /note= "Lectin domain"
FT Domain 160..193
FT /note= "egf domain"
FT 197..317
FT /note= "Complement factor binding domain"
FT 333..355
FT /note= "Transmembrane binding domain"
FT 356..372
FT /note= "Cytoplasmic domain"
XX
XX US525538-A.
XX
XX 06-JUL-1993.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 23-FEB-1989; 89US-0315015.
XX 22-NOV-1989; 89US-0440625.
XX 16-DEC-1991; 91US-0808122.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX WPI; 1993-226664/28.
XX N-PSDB; AAQ44243.
XX
XX New lymphocyte homing receptor immunoglobulin fusion
XX polypeptide(s) - used to inhibit binding of lymphocytes in
XX therapeutic and diagnostic uses
XX
XX Disclosure; Fig 1; 44pp; English.
XX
XX The sequences given in AAR38908-09 represent human and murine lymphocyte
XX cell surface glycoprotein (LHR) respectively. These proteins mediate
XX the binding of lymphocytes to the endothelium of lymphoid tissue. LHR
XX is a glycoprotein which contains a signal domain, a carbohydrate
XX binding domain, an epidermal growth factor-like (egf) domain, at least
XX one complement binding domain repeat, a transmembrane binding domain
XX (TMD) and a charged intracellular or cytoplasmic domain. The murine
XX (83%), however degrees of homology between the various domains is
XX variable. These proteins may be fused to a ligand binding partner
XX protein (LBPP) which causes an increase in the half life of the LHR.
XX The fusions may be used therapeutically to compete with the normal
XX binding of lymphocytes to lymphoid tissue. They may be used in organ
XX or graft rejection and for the treatment of inflammation.
XX
XX Sequence 372 AA:
SQ

```

```

Query Match 100.0%; Score 666; DB 14; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 WTYHYSEKPMNMOARARFCNDYTDLVAIQKAEIETLEKTLPEFSRSYWGIRKIGIW 60
DB 39 WTYHYSEKPMNMOARARFCNDYTDLVAIQKAEIETLEKTLPEFSRSYWGIRKIGIW 98
QY 61 TWVGTNKSLEBAENMGDGPNNKKKNEKDCVEIYIRKNDAGKWNDACHKRLKALIC 117
DB 99 TWVGTNKSLEBAENMGDGPNNKKKNEKDCVEIYIRKNDAGKWNDACHKRLKALIC 155

```

```

RESULT 21
AAR76506
ID AAR76506 standard; Protein; 372 AA.
XX
XX AAR76506;
XX

```

```

DT 01-DEC-1995 (first entry)
XX
XX Human LHR.
XX
XX Lymphocyte homing receptor; lymphocyte cell surface glycoprotein;
XX LHR; ligand binding partner; immunoglobulin; constant region;
XX antibody engineering; immunomodulator.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 20..32
XX /label= Sig.peptide
XX /note= "hydrophobic domain, may act as signal
XX for insertion into the endoplasmic
XX reticulum lumen"
XX
XX Protein 39..372
XX label= Mat_protein
XX 39..334
XX /note= "putative extracellular domain"
XX
XX Modified-site 60..62
XX /label= N-glycosylation-site
XX 104..106
XX /label= N-glycosylation-site
XX 177..179
XX /label= N-glycosylation-site
XX 216..218
XX /label= N-glycosylation-site
XX 232..234
XX /label= N-glycosylation-site
XX 246..248
XX /label= N-glycosylation-site
XX 271..273
XX /label= N-glycosylation-site
XX 311..313
XX /label= N-glycosylation-site
XX 335..357
XX /note= "putative stop transfer or membrane anchor
XX domain"
XX 358..372
XX /note= "putative intracellular region"
XX
XX Region
XX
XX US5428130-A.
XX
XX 27-JUN-1995.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 16-DEC-1991; 91US-0808122.
XX 08-DEC-1992; 92US-0986931.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX WPI; 1995-240086/31.
XX N-PSDB; AAQ92802.
XX
XX New hybrid ligand binding partner molecules - fused to immunoglobulin
XX constant region sequences to increase stability and in vivo plasma
XX half-life
XX
XX Disclosure; Fig.1a-1c; 40pp; English.
XX
XX A murine Mel 14 antigen cDNA clone was used to screen a lambda gt10
XX cDNA library derived from human peripheral blood lymphocyte mRNA
XX obtd. from primary cells. A cDNA clone encoding LHR was isolated.
XX
XX Sequence 372 AA:
SQ

```

Query Match 100.0%; Score 666; DB 16; Length 372;
 Best Local Similarity 100.0%; Pred. No. 2,3e-63;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRRRCRDNTDLVAIQNKAETLEYLEKTLPEFSRSYWGIRKIGIW 60
 |||
 Db 39 WYHYSEKPMNQRRRCRDNTDLVAIQNKAETLEYLEKTLPEFSRSYWGIRKIGIW 98

OY 61 TWVGTNKSITTEAEENMGDEPNKKKKEPCVEIYIKRNKDAGKWNDDACHKIKAAIC 117
 |||
 Db 99 TWVGTNKSITTEAEENMGDEPNKKKKEPCVEIYIKRNKDAGKWNDDACHKIKAAIC 155

RESULT 22
 AAR83050
 ID AAR83050 standard; Protein; 372 AA.
 AC AAR83050;
 XX
 DT 31-JAN-1996 (first entry)
 XX
 DE Human LHR`
 XX
 KW Lymphocyte cell surface glycoprotein; LHR; transmembrane receptor;
 KW Immunoglobulin; Igg; constant region; receptor-mediated disease;
 KW vector; plasma-life.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 20..32
 FT /label= Sig_peptide
 FT 39..155
 FT /label= Carbohydrate_binding_domain
 FT 160..193
 FT /label= Epidermal_growth_factor_domain
 FT 197..317
 FT /label= Complement_factor_binding_domain
 FT 333..355
 FT /label= Transmembrane_binding_domain
 FT 356..372
 FT /label= Cytoplasmic_domain
 FT Domain
 FT Modified-site 60..62
 FT /label= N-glycosylation_site
 FT Modified-site 104..106
 FT /label= N-glycosylation_site
 FT Modified-site 177..179
 FT /label= N-glycosylation_site
 FT Modified-site 216..218
 FT /label= N-glycosylation_site
 FT Modified-site 232..234
 FT /label= N-glycosylation_site
 FT Modified-site 248..248
 FT /label= N-glycosylation_site
 FT Modified-site 271..273
 FT /label= N-glycosylation_site
 FT Modified-site 311..313
 FT /label= N-glycosylation_site
 XX
 PN US5455165-A.
 XX
 PD 03-OCT-1995.
 XX
 PF 23-FEB-1989; 89US-0315015.
 XX
 PR 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185669.
 XX
 PA (GETH) GENENTECH INC.

XX
 PI Capon DJ, Laaky LA;
 XX
 DR WPI: 1995-350776/45.
 DR N-PSDB: AAT05869.
 PT Expression vector encoding fusion protein to increase plasma life -
 PT compresses receptor ligand binding site and Ig constant region, for
 PT treatment of receptor mediated disease
 XX
 PS Disclosure: Fig 1; 42pp; English.
 XX
 CC A mouse LHR (Lymphocyte cell surface glycoprotein) cDNA clone was used
 CC to screen an oligo-dT primed lambda gt10 cDNA library derived from
 CC human peripheral blood lymphocyte mRNA obtd. from primary cells. A
 CC 2.2 kb clone (sequence given in AAT05869) was isolated that encoded the
 CC human LHR protein (AAR83050). LHR-IgG hybrids were constructed for use
 CC in the targeting of therapeutic moieties to lymphoid tissue.
 SQ Sequence 372 AA;

Query Match 100.0%; Score 666; DB 16; Length 372;
 Best Local Similarity 100.0%; Pred. No. 2,3e-63;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRRRCRDNTDLVAIQNKAETLEYLEKTLPEFSRSYWGIRKIGIW 60
 |||
 Db 39 WYHYSEKPMNQRRRCRDNTDLVAIQNKAETLEYLEKTLPEFSRSYWGIRKIGIW 98

OY 61 TWVGTNKSITTEAEENMGDEPNKKKKEPCVEIYIKRNKDAGKWNDDACHKIKAAIC 117
 |||
 Db 99 TWVGTNKSITTEAEENMGDEPNKKKKEPCVEIYIKRNKDAGKWNDDACHKIKAAIC 155

RESULT 23
 AAR98106
 ID AAR98106 standard; Protein; 372 AA.
 AC AAR98106;
 XX
 DT 31-OCT-1996 (first entry)
 XX
 DE Human lymphocyte cell surface glycoprotein (HuLHR).
 XX
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..38
 FT /label= Signal region.
 FT 39..155
 FT /label= Lectin domain.
 FT 160..193
 FT /label= EGF domain.
 FT 197..258
 FT /label= Complement binding repeat 1.
 FT 259..317
 FT /label= Complement binding repeat 2.
 FT 333..355
 FT /label= Transmembrane domain.
 FT 356..372
 FT /label= Cytoplasmic domain.
 XX
 PN US5514582-A.
 XX
 PD 07-MAY-1996.
 XX
 PF 23-FEB-1989; 89US-0315015.

XX 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 XX WPI; 1996-238773/24.
 DR N-PSDB; AAR98106.
 DR
 XX
 Nucleic acid encoding hybrid immunoglobulin comprising the ligand
 binding site of a receptor fused to Ig constant region - useful for
 diagnosis and treatment e.g. of inflammation
 PT
 PR
 PT
 XX
 Example 2; Figure 1; 41pp; English.
 PS

CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region: fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation, metastasis of lymphoma etc.).
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes.
XX
XX Sequence 372 AA:

Query Match	100.0%;	Score 666;	DB 17;	Length 372;
Best Local Similarity	100.0%;	Pred. No. 2.3e-63;		
Matches 117;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	WT	HY	SE	KP	MM	Q	R	A	R	R	F	C	R	D	N	T	D	V	A	I	O	N	K	A	E	I	E	L	E	K	T	P	S	R	S	Y	Y	W	I	G	I	R	K	I	G	I	W	60					
Db	39	W	T	H	Y	S	E	K	P	M	M	Q	R	A	R	R	F	C	R	D	N	T	D	V	A	I	O	N	K	A	E	I	E	L	E	K	T	P	S	R	S	Y	Y	W	I	G	I	R	K	I	G	I	W	98

Dy 61 TWVGTKSLTEEAENWGDGEPANNNKKECEVEIYIRNKDAGKNDDACHIKALC 117
|||
99 twvgtnkslteeenwgdgpnknkkelceveiyirkndagkwddachiklaalc 155

RESULT 24

AA98133
ID AA98133 standard; Protein; 372 AA.

AC AAR98133;

DT 01-NOV-1996 (first entry)

Human lymphocyte cell surface glycoprotein (HULHR) variant.

Immunoglobulin; transmembrane receptor; adhesion; targeting; diagnosis; therapy; drug delivery; antiviral; neuromodulator; immunomodulator; cell adhesion; graft rejection; inflammation; metastasis.

05 Homo sapiens.

FH	key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

2

FT	Region	1..37
FT	/label= signal region.	
FT	Domain	38..154
FT	/label= Lectin domain.	
FT	Domain	159..192
FT	/label= EGF domain.	
FT	Binding-site	196..258
FT	/label= Complement binding repeat 1	
FT	Binding-site	259..317
FT	/label= Complement binding repeat 2	
FT	Domain	333..355
FT	/label= Transmembrane domain.	
FT	Domain	356..372
FT	/label= Cytoplasmic domain.	

PN	US514582-A.	
XX		
PD	07-MAY-1996.	
XX		
PE	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.

PA (GETH) GENENTECH INC.

PI Capon DJ, Lasky LA;

DR WPI; 1996-238773/24.

PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation
----	--

PS Disclosure; Page 19; 41pp; English

A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region; fused at its C-terminus with the N-terminus of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targeting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the *in vitro* assay of LBP and their targets; or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particular cells. Typical applications are as antiviral, neuromodulating and immunomodulating agents, or as modulators of cell adhesion (e.g. in treating graft rejection; inflammation; metastasis of lymphoma etc.). The immunoglobulin component increases plasma half life and facilitates purification while detection of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR8106 are given in AAR8109-RB8135. This variant contains a 209-Asn insertion.

Sequence 372 AA;

Query Match	100.0%;	Score 666;	DB 17;	Length 372;
Best Local Similarity	100.0%;	Pred. No. 2.3e-63;		
Matches 117;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

1 WTYHYSEKPMNWQRARRECRDNYTDLVAIQNKAEIEYLEKTLPESSRYWIGIRKIGIW 60

38 wlynysekpmmwqrarrtcrdnytdlvaigkaeieylektlpsrsywigirkigiw

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

SO	Sequence	372 AA;
Query Match	100.0%;	Score 666; DB 22; Length 372;
Best Local Similarity	100.0%	
DNA W 3	2	

	Matches	117;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
Oy	1	WTHYSEKPNMORARFCRDNYTDLVAIONKAETLEYLEKTLPFSRSYWGIGRTGICW	60							
bh	39	wtbhvskkkmnarrararfndntdlunlrbhl	60							

61 TWGTGKSLTEELDENMGDEEPNNKKNKEDCEYIIYIKRKDKAGKWNDDACHKRLKALC 117
99 twgtgkslteeenmgddennknkknkdcceiylkrrkdkdgvwnddachrlkallc 155			

28 RESULT
28 AAR98134

RESULT 28
AAR98134
ID AAR98134 standard; Protein; 374 AA.
XX
AC AAR98134:

ID	AA98134	standard; Protein; 374 AA.
XX		
AC	AA98134;	
XX		
DT	01-NOV-1996	(first entry)

AC	AAR98134;
XX	
DT	01-NOV-19
XX	
DE	Human 1 ym

DT	01-NOV-1996	(first entry)
XX		
DE	Human lymphocyte cell surf	
XX		
KW	Immunoglobulin: transmembr	

DE	Human lymphocyte cell surface glycoprotein (HLADR) variant
XX	
KW	Immunoglobulin; transmembrane receptor; adhesion; targetting
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator; cell
KW	immunomodulator; cell adhesion; graft rejection; inflammation

KW	immunoglobulin; transmembrane receptor; adhesion; targeting;
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW	immunomodulator; cell adhesion; graft rejection; inflammation;
KW	metastasis.
XX	

KW	immunomodulat
KW	metastasis.
XX	
OS	Homo sapiens
XX	

	Location/Qualifiers
OS	Homo sapiens.
XX	
FH	Key
FT	Region
FT	1..37
	/label=Stimul root

Key	Location/Qualifiers
FH	1..37
FT	/label= Signal region
FT	38..154
FT	/label= Tectin domain

FT	/label=	Signal region
FT	38.154	
FT	/label=	Lectin domain
FT	159.192	
FT	/label=	EGF domain
FT		

FT	/label= Lectin domain
FT	159..192
FT	/label= EGF domain.
FT	196..260
FT	/label= Complement binding site

FT	/label= EGF domain.
Binding-site	196..260
FT	/label= Complement binding repeat 1
Binding-site	261..319
FT	/label= Complement binding repeat 2

FT	Binding-site	Domain
FT	1	1
FT	2	2
FT	3	3
FT	4	4
FT	5	5
FT	6	6
FT	7	7
FT	8	8
FT	9	9
FT	10	10
FT	11	11
FT	12	12
FT	13	13
FT	14	14
FT	15	15
FT	16	16
FT	17	17
FT	18	18
FT	19	19
FT	20	20
FT	21	21
FT	22	22
FT	23	23
FT	24	24
FT	25	25
FT	26	26
FT	27	27
FT	28	28
FT	29	29
FT	30	30
FT	31	31
FT	32	32
FT	33	33
FT	34	34
FT	35	35
FT	36	36
FT	37	37
FT	38	38
FT	39	39
FT	40	40
FT	41	41
FT	42	42
FT	43	43
FT	44	44
FT	45	45
FT	46	46
FT	47	47
FT	48	48
FT	49	49
FT	50	50
FT	51	51
FT	52	52
FT	53	53
FT	54	54
FT	55	55
FT	56	56
FT	57	57
FT	58	58
FT	59	59
FT	60	60
FT	61	61
FT	62	62
FT	63	63
FT	64	64
FT	65	65
FT	66	66
FT	67	67
FT	68	68
FT	69	69
FT	70	70
FT	71	71
FT	72	72
FT	73	73
FT	74	74
FT	75	75
FT	76	76
FT	77	77
FT	78	78
FT	79	79
FT	80	80
FT	81	81
FT	82	82
FT	83	83
FT	84	84
FT	85	85
FT	86	86
FT	87	87
FT	88	88
FT	89	89
FT	90	90
FT	91	91
FT	92	92
FT	93	93
FT	94	94
FT	95	95
FT	96	96
FT	97	97
FT	98	98
FT	99	99
FT	100	100

FT	/label= Complement binding
FT	335..357
FT	/label= Transmembrane domain
FT	358..374
FT	/label= Cytoplasmic domain

	Domain
FT	
FT	
FT	
XX	
PN	US5514582-A.

FT	
XX	US5514582-A.
PN	
XX	07-MAY-1996.
PD	

PN	US5514582-A.
XX	
PD	07-MAY-1996.
XX	
PF	23-FEB-1989.

PD	0/-MAY-1996.	
XX		
PF	23-FEB-1989;	89US-0315015,
XX		
PR	22-NOV-1989;	89US-0440625,

PR	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.

PR 23-FEB-1989; 89US-0315015
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.

PR	16-DEC-1991;	910S-08
PR	08-DEC-1992;	920S-09
PR	21-JAN-1994;	940S-01
XX		
PA	(GETH) GENENTECH INC.	

PR 21-JAN-1994; 9405
XX
PA (GETH) GENENTECH INC
XX
PI Capon DJ, Lasky LA

PA (GEIH) GENENIECH I
XX
PI Capon DJ, Lasky LA,
XX
DRI WPI; 1996-238773/24

Capon DJ, Lasky LA;
WPI, 1996-238773/24.
Nucleic acid encoding hybrid immunoglobulin comprising the ligand

Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation

Disclosure, Page 19; 41pp; English.


```

FH Key Location/Qualifiers
FT Region 347..368
FT /label= transmembrane
FT Modified-site 113..115
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 157..159
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 230..232
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 269..271
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 285..287
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 299..301
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 324..326
FT /label= N-linked_glycosylation
FT /note= "putative"
XX
XX W09201049-A.
XX
XX 23-JAN-1992.
XX
XX 15-JUL-1990; 90WO-US04986.
XX
XX 13-JUL-1990; 90US-0553759.
XX
XX (GEHO-) GEN HOSPITAL CORP.
XX
XX Seed B, Aruffo A, Amlot M;
XX
XX WPI; 1992-056864/07.
XX
XX N-PSDB; AAQ21184.
XX
XX New CD53 cell surface antigen and DNA encoding it - for
XX immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
XX Example 14; Page 106; 160pp; English.
XX
XX Two CDNA clones encoding Leu8 determinants were isolated from a
XX human T cell library using the rapid immunoselection cloning method
XX (see e.g. AAQ21164 for description of method). This protein sequence
XX was deduced from the larger insert. The hydrophobic putative
XX membrane-spanning domain is followed by several positively charged
XX residues resembling a cytoplasmic anchor sequence. The protein
XX is closely related to the murine Mel-14 homing receptor.
XX See AAQ25500 for the shorter insert and AAR22551 for the phospholipid
XX anchored form of Leu8 antigen.
XX
XX Sequence 385 AA:

```

```

Query Match 100.0%; Score 666; DB 13; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,4e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 WTHYSEKPNWRRARFCNDNTDVAIONKAEIYLEKTPFSRSYWGIRKIGIWM 60
OY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 Wcyhysekpmnwgrarrfcndytdlvalqnkaeleylktlprfsrwygirkigw 111
OY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 61 TWVGTNKSITLFEAEWNGDGEPPNKKKKEPCVEYIYIRKNDAGKWNDDACHKKAALC 117
OY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 112 twvgtnkslteeenwgdgeppnknkcdcevelyikrnkdagkwndachkikaalc 168

```

```

RESULT 31
AAR32707

```

```

ID AAR32707 standard; Protein: 385 AA.
XX
XX AAR32707;
AC
XX 14-JUN-1993 (first entry)
XX
XX LAM-1 from pLAM-1.
XX
XX Leukocyte adhesion molecule; epitope; monoclonal antibody;
XX inflammation; autoimmune disease; shock; tissue damage; organ;
XX tissue; transplant; metastasis.
XX Homo sapiens.
XX
XX Location/Qualifiers
FH Key 13..41
FH Region /note= "hydrophobic, possible signal sequence"
FT Region 346..362
FT /note= "hydrophobic, possible transmembrane sequence"
FT Modified-site 73..75
FT /note= "N-glycosylation site"
FT Modified-site 117..119
FT /note= "N-glycosylation site"
FT Modified-site 117..119
FT /note= "N-glycosylation site"
FT Modified-site 190..192
FT /note= "N-glycosylation site"
FT Modified-site 245..247
FT /note= "N-glycosylation site"
FT Modified-site 259..261
FT /note= "N-glycosylation site"
FT Modified-site 284..286
FT /note= "N-glycosylation site"
FT Modified-site 324..326
FT /note= "N-glycosylation site"
XX
XX W09302698-A.
XX
XX 18-FEB-1993.
XX
XX 23-JUL-1992; 92WO-US06127.
XX
XX 29-JUL-1991; 91US-0737092.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Spertini OG, Tedder TF;
XX
XX WPI; 1993-076177/09.
XX
XX N-PSDB; AAQ37304.
XX
XX Monoclonal antibodies to leukocyte adhesion molecule-1 (LAM-1) -
XX useful for treating inflammation, cancer, auto-immune diseases,
XX tissue damage and organ or tissue transplant rejection
XX
XX Disclosure; Fig 2; 121pp; English.
XX
XX B cell-specific cDNA was isolated from a human tonsil cDNA library
XX (ATCC #37546) using differential hybridisation with labelled cDNAs
XX from either B cell(RAJ1) RNA or T cell (HSB-2) RNA. Positive
XX plaques were isolated and cloned and the cDNA inserts subcloned into
XX pSP65. One of the 261 RAJ1+ HSB2- cDNA clones isolated, B125,
XX contained a 1.9 kb cDNA insert that hybridised with a 2.4 kb RNA
XX species found in several B cell lines. However B125 did not
XX hybridise with any of the other RAJ1+ HSB- clones or with mRNA from
XX several T cell lines. The B125 cDNA clone hybridised with a 2.3 kb
XX cDNA, termed pLAM-1 (prod. shown). The amino acid sequence of
XX LAM-1 indicates a structure typical of a membrane glycoprotein.
XX Monoclonal antibodies specific for LAM-1 may be prep. and are useful
XX in the treatment of inflammation, autoimmune disease, shock, tissue
XX damage, organ or tissue transplant rejection and to inhibit the
XX metastasis and homing of malignant cells which express the LAM-1
XX receptor protein. They can be used to monitor lamy-1 gene regulation,
XX

```


CC e.g. in drug or therapy tests.
XX
SQ Sequence 385 AA:

Query Match 100.0%; Score 666; DB 14; Length 385;
Best Local Similarity 100.0%; Pred. No. 2.4e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WYHYSEKPMNQARFRFRDNTDVAIQNKAIEYLEKTLPEFSRSTYWGIRKIGTW 60
|
DB 52 WYHYSEKPMNQARFRFRDNTDVAIQNKAIEYLEKTLPEFSRSTYWGIRKIGTW 111
|

OY 61 TWVGTKSLTEAEKNMGDEPNKKKEDCVELIYIRNKDAGKWNDDACHKLKALC 117
|
DB 112 twvgtnkslteaeenwgdgpnkknkdcveliykrnkdagkwndachklkalc 168
|

RESULT 32
AAR34197
ID AAR34197 standard; Protein: 385 AA.
XX
AC AAR34197;
XX
DT 17-AUG-1993 (first entry)
XX
DE Sequence encoded by LAM-1 cDNA.
XX
KW Leukocyte adhesion molecule-1; lymphocyte-associated molecule-1;
KM LAM-1.
XX
OS Homo sapiens.
XX

EH Key Location/Qualifiers
FT Binding-site 73..75
FT Binding-site /label= N-linked glycosylation
FT Binding-site 117..119
FT Binding-site /label= N-linked glycosylation
FT Binding-site 190..192
FT Binding-site /label= N-linked glycosylation
FT Binding-site 245..247
FT Binding-site /label= N-linked glycosylation
FT Binding-site 259..261
FT Binding-site /label= N-linked glycosylation
FT Binding-site 284..286
FT Binding-site /label= N-linked glycosylation
FT Binding-site 324..326
FT Binding-site /label= N-linked glycosylation
FT Region 13..41
FT Region /label= Hydrophobic
FT Region 346..362
FT Region /label= Hydrophobic
FT

PN WO9306835-A.
XX
XX
PD 15-APR-1993.
XX
XX
PF 05-OCT-1992; 92WO-2008467.
XX
XX
PR 03-OCT-1991; 91US-0770608.
XX
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
XX
PI Tedder TF, Spertini OG;
XX
XX
DR WPI: 1993-134119/16.
XX
DR N-PSDB: AAO38839.
XX

PT Detection of inflammation or disease site - by using labelled
PT Leukocyte adhesion molecule-1 - treats leukocyte-mobilising
PT conditions, auto-immune disorders, cancer etc.
XX
XX
PS Disclosure: Fig 2; 44pp; English.

XX
CC cDNA encoding the LAM-1 protein was isolated from a human tonsil
CC CDNA library (ATCC 37546). The mature LAM-1 protein has an
CC extracellular region of about 294 amino acids containing 7 potential
CC N-linked carbohydrate attachment sites. LAM-1 has a cytoplasmic tail
CC of 17 amino acids containing 8 basic and 1 acidic residues. The
CC processed LAM-1 protein has a Mr of at least 50,000. LAM-1 combines
CC domains homologous to domains found in three distinct families of
CC molecules: animal lectins, growth factors, and C3/C4 binding proteins.
XX

SQ Sequence 385 AA:

Query Match 100.0%; Score 666; DB 14; Length 385;
Best Local Similarity 100.0%; Pred. No. 2.4e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WYHYSEKPMNQARFRFRDNTDVAIQNKAIEYLEKTLPEFSRSTYWGIRKIGTW 60
|
DB 52 WYHYSEKPMNQARFRFRDNTDVAIQNKAIEYLEKTLPEFSRSTYWGIRKIGTW 111
|

OY 61 TWVGTKSLTEAEKNMGDEPNKKKEDCVELIYIRNKDAGKWNDDACHKLKALC 117
|
DB 112 twvgtnkslteaeenwgdgpnkknkdcveliykrnkdagkwndachklkalc 168
|

RESULT 33
AAR56663
ID AAR56663 standard; Protein: 385 AA.
XX
AC AAR56663;
XX
DT 25-MAR-1995 (first entry)
XX
DE L-selectin.
XX
DE L-selectin.
XX
KW L-selectin; lectin domain; bifunctional; ligand-binding polypeptide;
KW Leucocyte; mobilising activity; tissue damage; autoimmune disease;
KW post-reperfusion injury; organ/tissue transplant rejection;
KW inflammation; inhibition; platelet; mononuclear cell; aggregation;
KW thrombosis.
XX
XX
OS Homo sapiens.
XX
XX
PN WO9417193-A.
XX
XX
PD 04-AUG-1994.
XX
XX
PF 25-JAN-1994; 94WO-US00909.
XX
XX
PR 25-JAN-1993; 93US-0008459.
XX
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
XX
PI Kansas GS, Tedder TF;
XX
XX
DR WPI: 1994-264106/32.
XX
DR N-PSDB: AAO71006.
XX

PT New bifunctional polypeptide contg. binding domains of different
PT selectin - and related DNA useful as leucocyte mobilising agents
PT for treating autoimmune disease, post-reperfusion injury etc.,
PT also for targeted drug delivery.
XX
XX
PS Disclosure: Page 26-28; 56pp; English.
XX

This sequence represents L-selectin. The lectin domain of L-selectin
may be used in the construction of the new bifunctional ligand-
binding polypeptide of the invention. These peptides have leucocyte
mobilising activity for treating tissue damage, autoimmune disease,
post-reperfusion injury or organ/tissue transplant rejection. They
act by interfering with binding of leucocytes at the site of
inflammation and also inhibit platelet and mononuclear cell

CC aggregation so prevent thrombosis. These bifunctional peptides can
CC bind to 2 different selectin ligands so provide more effective
CC targeting than monospecific compounds.
XX
SO Sequence 385 AA:

Query Match 100.0%; Score 666; DB 15; Length 385;
Best Local Similarity 100.0%; Pred. No. 2.4e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WTYHYSEKPMNQRARRECRDNYTDVAIONKAEIEYEKLTLPFSRSTYWGIRKIGTM 60
Db |||||||
52 WTYHYSEKPMNQRARRECRDNYTDVAIONKAEIEYEKLTLPFSRSTYWGIRKIGTM 111
QY 61 TWVGTKNSLTTEAENMGDEPNKKKKEDCVETIYIRKNDAGKWNDDACHKILKAALC 117
Db |||||||
112 twvgtnkslteaenwgdgpnknkknkedcveiyikrnkdagkwndachkilaalc 168

RESULT 34

AAR91442
ID AAR91442 standard; Protein: 385 AA.

AC AAR91442;

DT 31-OCT-1996 (first entry)

DE Human Leu8 antigen.

XX Cell surface antigen; cloning; immunoselection; immunotherapy;

KW therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;

XX antiinflammatory.

OS Homo sapiens.

FT Key Location/Qualifiers
FT Modified-site 113..115 /label= glycosylation_site
FT Modified-site 156..158 /label= glycosylation_site
FT Modified-site 229..231 /label= glycosylation_site
FT Modified-site 269..271 /label= glycosylation_site
FT Modified-site 285..287 /label= glycosylation_site
FT Modified-site 299..301 /label= glycosylation_site
FT Modified-site 324..326 /label= glycosylation_site
FT Domain /label= glycosylation_site
FT 347..368 /label= Transmembrane_domain

PN USS506126-A.

PD 09-APR-1996.

PF 25-FEB-1988; 88US-0160416.

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 18-OCT-1993; 93US-0139273.

PA (GEHO) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

XX WPI; 1996-200279/20.

DR N-PSDB; AAT14723.

PT Cloning of cDNA encoding cell surface antigen - useful for isolation
PT of diagnostic and therapeutic proteins
XX
XX Example 14; Column 71-74; 79pp; English.
CC 2 cDNA clones encoding Leu8 determinants (AAR91442 and AAR91443) were
CC isolated from a human T-cell library using a novel immunoselection
CC cloning method. The longer insert (AAT14723) contained 2,350
CC residues, while the shorter lacked 436 internal residues. A major
CC transcript of 2.4 kb was present in peripheral blood mononuclear
CC cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,
CC and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukaemic
CC T-cell lines. Leu8 antigens can be obtd. for diagnostic and
CC therapeutic use. The presence or absence of Leu8 on CD4+ T-cells
CC identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.
CC Soluble forms of Leu8 can act as antiinflammatory agents by reducing
CC lymphocyte migration.
XX
SO Sequence 385 AA:

Query Match 100.0%; Score 666; DB 17; Length 385;
Best Local Similarity 100.0%; Pred. No. 2.4e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQRARRECRDNYTDVAIONKAEIEYEKLTLPFSRSTYWGIRKIGTM 60
Db |||||||
52 WTYHYSEKPMNQRARRECRDNYTDVAIONKAEIEYEKLTLPFSRSTYWGIRKIGTM 111
QY 61 TWVGTKNSLTTEAENMGDEPNKKKKEDCVETIYIRKNDAGKWNDDACHKILKAALC 117
Db |||||||
112 twvgtnkslteaenwgdgpnknkknkedcveiyikrnkdagkwndachkilaalc 168

RESULT 35

AAW21657
ID AAW21657 standard; Protein: 385 AA.

AC AAW21657;

DT 30-SEP-1997 (first entry)

DE Human lymphocyte-associated cell surface protein.

XX Lymphocyte-associated cell surface protein; LAM-1; B lymphocyte;

KW antibody; autoimmune disease; cancer.

XX Homo sapiens.

FT Key Location/Qualifiers
FT Peptide 1..51 /label= Sig-peptide
FT Peptide 14..51 /label= Sig-peptide
FT /note= "Signal peptide starting from alternative
FT N-terminal Met"
FT Protein 52..385 /label= Mat_protein
FT 52..345 /label= Extracellular_domain
FT 346..362 /label= Transmembrane_domain
FT 363..385 /label= Intracellular_domain
FT Modified-site 73..75 /label= glycosylation
FT /note= "putative N-glycosylation site"
FT 117..119 /label= glycosylation
FT /note= "putative N-glycosylation site"
FT 190..192 /label= glycosylation
FT /note= "putative N-glycosylation site"

FT	Modified-site	245..247	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
FT	Modified-site	259..261	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
FT	Modified-site	288..290	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
FT	Modified-site	324..326	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
XX			
PN	EP770680-A2.		
XX			
PD	02-MAY-1997.		
XX			
PF	21-FEB-1990;	90EP-030184.	
XX			
PR	21-FEB-1989;	89US-0313109.	
XX			
PA	(DAND) DANA FARBER CANCER INST INC.		
XX			
PI	Tedder TF:		
XX			
DR	WPI: 1997-238140/22.		
DR	N-PSDB: AAT72270.		
XX			
PT	New antibodies to leukocyte adhesion molecule-1 - used for binding		
PT	LAM-1, e.g. for treating tissue damage, autoimmune disorders,		
PT	cancers or organ or tissue transplants		
XX			
PS	Claim 1; Page 6-8; 10pp: English.		
XX			
CC	Human lymphocyte-associated cell surface protein LAM-1 (AA021657)		
CC	contains domains homologous with binding domains of animal lectins,		
CC	growth factors and C3/C4 binding proteins. Its amino acid sequence		
CC	was deduced from a human tonsil cDNA clone (AAT72270). Together with		
CC	ELAM-1 and GMP-40, LAM-1 defines a new family of homologous		
CC	structures that are expressed by different cell lineages and which		
CC	can function as receptors in cellular interactions. LAM-1 can be		
CC	used as an antigen to raise antibodies capable of preventing		
CC	adhesion, migration or infiltration into tissues of cells		
CC	expressing the protein, or to identify antagonists useful for the		
CC	treatment of tissue damage, autoimmune disorders, cancer, or organ		
CC	or tissue transplants.		
XX			
XX	Sequence 385 AA:		
SO			
Query Match 100.0%; Score 666; DB 18; Length 385;			
Best Local Similarity 100.0%; Pred. No. 2.4e-63;			
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 WTYHYSEKPMWQARRCRDNYTDLVAIQNKAIELEYLEKTLPSFSRYWIGIRKIGTW 60		
DB	52 WCYHYSEKPMWQARRCRDNYTDLVAIQNKAIELEYLEKTLPSFSRYWIGIRKIGTW 111		
QY	61 TWVGNGSKLTLEAEWNGGSEPRNNKKNKEDCVETIYIKRNKDAGKWDACHTKKAALC 117		
DB	112 TWVGNGSKLTLEAEWNGGSEPRNNKKNKEDCVETIYIKRNKDAGKWDACHTKKAALC 168		
RESULT 36			
ID	AA080452		
XX	AA080452 standard; Protein; 385 AA.		
XX	AA080452;		
XX	07-JUN-1999 (first entry)		
XX	Human Leu8 antigen.		
XX			

Key	Location/Qualifiers
Domain	347..368
FT	/note="transmembrane domain"
XX	US5830731-A.
XX	03-NOV-1998.
XX	21-MAY-1997; 9705-0861205.
XX	01-DEC-1992; 9205-0983647.
XX	25-FEB-1988; 8805-0160416.
PR	13-JUL-1989; 8905-0379076.
PR	13-JUL-1990; 9005-0553759.
PR	21-MAY-1997; 9705-0861205.
PA	(GEHO) GEN. HOSPITAL CORP. *
PI	Aruffo A, Seed B:
XX	WPI: 1998-609251/51.
DR	N-PSDB; AAV63460.
XX	New cloning vector and polylinker - based on existing sequences for efficient cloning and expression of mammalian cDNA(s), especially human lymphocyte antigenic sequences
PT	Example 14; Column 67-70; 75pp: English.
PS	This polypeptide comprises human Leu8 antigen. Its amino acid sequence was deduced from the nucleotide sequence (see AAV63450) of a cDNA clone isolated from a T lymphocyte cDNA library using a novel method for cloning cDNAs from mammalian expression libraries.
XX	The method is based on transient expression of an antigen in eukaryotic cells and physical selection of cells expressing the antigen by adhesion to an antibody-coated substrate. It is useful for the isolation and cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell, and was used to clone genes (see AAV63442-63) encoding cell surface antigens from mammalian lymphocytes (see AAV80440-55). The purified genes and proteins are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders of animals, including humans. The extracellular domain of Leu8, which mediates adhesion to specialised endothelial cells of lymph nodes, is quite specific in its recognition of the lectin ligand sulphated galactosyl ceramide (sulfatide). Modification of the specificity of this binding could serve to regulate the homing potential of resting T cells. Soluble forms of Leu8 can act as antiinflammatory agents by reducing lymphocyte migration.
CC	Sequence 385 AA:
CC	Query Match 100.0%; Score 666; DB 19; Length 385;
CC	Best Local Similarity 100.0%; Pred. No. 2,4e-63;
CC	Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC	1 WYHYHSEKPMNQARRRFRDNYTDLVAIQNAELFELEKTIPIFSRSYTWIGIRKIGTW 60
DB	52 WYHYHSEKPMNQARRRFRDNYTDLVAIQNAELFELEKTIPIFSRSYTWIGIRKIGTW 111
QY	61 TWVGNKSLTEEAEMNGDESPNNKNKEDCVETIYIKRNKDACGMWDDACHRKAKALC 117
DB	112 TWVGNKSLTEEAEMNGDESPNNKNKEDCVETIYIKRNKDACGMWDDACHRKAKALC 168

AAW6199
 ID AAW6199 standard; Protein; 385 AA.
 AC AAW6199;
 DT 10-MAY-1999 (first entry)
 DE Human Leu8 antigen.
 KW Leu8; cell surface antigen; human; CDNA library; T lymphocyte; antiinflammatory; therapy.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 347..368
 FT /note="transmembrane domain"
 XX US5849898-A.
 XX 15-DEC-1998.
 XX 07-JUN-1995; 95US-0485447.
 XX 01-DEC-1992; 92US-0983647.
 XX 25-FEB-1988; 88US-0160416.
 XX 13-JUL-1989; 89US-0379076.
 XX 23-MAR-1990; 90US-0498809.
 XX 13-JUL-1990; 90US-0553759.
 XX 07-JUN-1995; 95US-0485447.
 PA (GENO) GEN HOSPITAL CORP.
 XX Allen J, Amiot M, Aruffo A, Cameroni D, Lauffer L;
 PI Oquendo C, Seed B, Simmons D, Stamenkovic I, Stengelin S;
 XX WPI: 1999-069813/06.
 DR N-PSDB; AAW61217.
 XX
 PT cDNA encoding human CD40 antigen - useful for cloning cDNA encoding
 PT cell surface antigens, constructing cDNA libraries, expression
 PT vectors for expression in eukaryotic cells or their fragments
 XX
 PS Example 14; Column 67-70; 79pp; English.
 XX
 CC This polypeptide comprises human Leu8 antigen. Its amino acid
 CC sequence was deduced from the nucleotide sequence (see AAW61217) of a
 CC cDNA clone isolated from a T lymphocyte cDNA library using a
 CC novel method for cloning cDNAs from mammalian expression libraries.
 CC The method is based on transient expression of an antigen in
 CC eukaryotic cells and physical selection of cells expressing the
 CC antigen by adhesion to an antibody-coated substrate. It is useful
 CC for the isolation and cloning of any protein which can be expressed
 CC and transported to the cell surface membrane of a eukaryotic cell,
 CC and was used to clone genes (see AAW81198-220) encoding cell surface
 CC antigens such as CD1a, CD1b, CD1c, CD2, CD6, CD7, CD13, CD14, CD16,
 CC CD19, CD20, CD22, CD26, CD27, CD28, CD31, CD32a, CD33b, CD33,
 CC CD34, CD36, CD37, CD38, CD39, CD40, CD43, CD44, CD53, ICAM, LFA-3,
 CC FCRII, FCRIIb, T118a and Leu8 (see AAW6188-62, AAW8151-52 and
 CC AAW88451). CD40 cDNA (see AAW81198) is specifically claimed. The
 CC extracellular domain of Leu8, which mediates adhesion to specialised
 CC endothelial cells of lymph nodes, is quite specific in its recognition of
 CC the lectin ligand sulphated galactosyl ceramide (sulfatide).
 CC Modification of the specificity of this binding could serve to
 CC regulate the homing potential of resting T cells. Soluble forms
 CC of Leu8 can act as antiinflammatory agents by reducing
 CC lymphocyte migration.
 CC
 XX Sequence 385 AA:

Query Match 100.0%; Score 666; DB 20; Length 385;
 Best Local Similarity 100.0%; Pred. No. 2.4e-63;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WTYHSEKPMNORARPCRDNYDVAIQNKATIELEKLPSSRSYTWIGIKIGIV 60
 DB 52 WCYHYSEKPMNMQARPCRDNYDVAIQNKATIELEKLPSSRSYTWIGIKIGIV 111
 QY 61 TWGVTNKSLEAEANWGDGEPNKKNEDECEIYIKRNKDGAKGNDDACRKAALC 117
 DB 112 TWGVTNKSLEAEANWGDGEPNKKNEDECEIYIKRNKDGAKGNDDACRKAALC 168
 RESULT 38
 ID AAY96138
 AC AAY96138 standard; Protein; 385 AA.
 DT 19-DEC-2000 (first entry)
 DE Human T-cell specific Leu8 antigen.
 KW Leu8; cell surface antigen; human; immunoselection; panning;
 KW immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; inflammation; antiinflammatory.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 347..368
 FT /label="Transmembrane_domain"
 XX US611093-A.
 XX 29-AUG-2000.
 XX 28-OCT-1998; 98US-0181612.
 XX 01-DEC-1992; 92US-0983647.
 XX 25-FEB-1988; 88US-0160416.
 XX 13-JUL-1989; 89US-0379076.
 XX 23-MAR-1990; 90US-0498809.
 XX 13-JUL-1990; 90US-0553759.
 PA (GENO) GEN HOSPITAL CORP.
 XX Stamenkovic I, Seed B;
 PI WPI: 2000-586382/55.
 DR N-PSDB; AAA50596.
 XX
 PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 PT useful for immunodiagnosis and immunotherapy of immune-mediated
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 PT diseases
 XX
 PS Example 14; Column 69-70; 75pp; English.
 XX
 CC The present sequence is that of a T-cell specific Leu8 antigen, as
 CC predicted from 1 of 2 clones (see AAA50596) isolated from a human
 CC T-cell library by the method of the invention. A shorter Leu8
 CC antigen is given in AAY96184. The method, designed to isolate cell
 CC surface antigen (CSA) cDNAs, is based upon transient expression of
 CC a CSA in eukaryotic cells and physical selection of cells expressing
 CC the antigen by adhesion to (panning on) an antibody-coated substrate
 CC such as a culture dish. CSA nucleic acids isolated by the method of
 CC the invention, and the proteins they encode, are useful for
 CC immunodiagnosis and immunotherapeutic applications, including the
 CC diagnosis and treatment of immune-mediated infections, diseases, and
 CC disorders in animals, including humans. These disorders include
 CC asthma, immune-complex disease, amyloidosis, parasitic diseases or
 CC multiple sclerosis. The ability to interfere with the binding of
 CC Leu8-T-cells to antigen presenting cells, or the ability to cause

Query Match	Best Local Similarity	100.0%;	Score 666;	DB 21;	Length 385;	
Matches 117;	Conservative	0;	Mismatches 0;	Indels	0;	Gaps 0;
OY	1	WTYHYSEKPMNQRRCRDNYTGLVAIVIONKAEIEYELEKTLTPFSRSYYWIGIRKIGIM	60			
DB	52	WLYHYSEKPMNQRRCRDNYTGLVAIVIONKAEIEYELEKTLTPFSRSYYWIGIRKIGIM	111			
OY	61	TWVGTSKSLTEBAENMGDEPNKKKNEKDCVEIYIKRNKADGKWNDDACHKLAALC	117			
DB	112	twvgtskslteeaenmgdepnkkknkedcveilylkrnkdgakwddachkikaalc	168			
RESULT	39					
AAU02447						
ID	AAU02447	standard;	Protein;	405	AA.	
XX	AAU02447;					
XX	29-AUG-2001	(first entry)				
DE	Human T-lymphocyte specific antigen Leu8 polypeptide.					
XX	Human; T-lymphocyte specific antigen; immune-mediated disease; Leu8;					
KW	infection; immune deficiency disorder; hypersensitivity; inflammation;					
KW	systemic lupus erythematosus; platelet disorder; rheumatoid arthritis;					
XX	transplant rejection; asthma.					
OS	Homo sapiens.					
XX						
XX	Key	Location/Qualifiers				
FT	Modified-site	113..115				
FT		/note= "N-linked glycosylation site"				
FT	Modified-site	157..159				
FT		/note= "N-linked glycosylation site"				
FT	Modified-site	230..232				
FT		/note= "N-linked glycosylation site"				
FT	Modified-site	269..271				
FT		/note= "N-linked glycosylation site"				
FT	Modified-site	285..287				
FT		/note= "N-linked glycosylation site"				
FT	Modified-site	324..326				
FT		/note= "N-linked glycosylation site"				
FT	Misc-difference	344..385				
FT		/note= "These amino acids are present in the longer form but are absent in the short form of Leu8"				
FT	Domain	347..368				
FT		/label= "transmembrane_domain"				
XX	US6218525-B1.					
XX	17-APR-2001.					
XX	01-DEC-1992;	92US-0983647.				
XX	25-FEB-1988;	88US-0160416.				
XX	13-JUL-1989;	89US-0379076.				
XX	13-JUL-1990;	90US-0553759.				
XX	(GEO) GEN HOSPITAL CORP.					

XX	Seed B, Aruffo A, Simmons D;
PI	WPI: 2001-289848/30.
DR	N-PsDB: AAS03190.
XX	
PT	New recombinant DNA encoding CD28 useful for diagnosing and treating
PT	immune-mediated diseases, infections or disorders; e.g. systemic lupus
PT	erythematosus, asthma, transplant rejection, rheumatoid arthritis
XX	
PS	Example 14; Column 63-66; 72pp; English.
CC	The present sequence represents human T-lymphocyte specific antigen
CC	Leub encoded by a long Leub cDNA clone. The shorter cDNA clone lacks 436
CC	bases and consequently lacks residues 344-385 given here. Various human
CC	lymphocyte cell surface antigen polypeptide sequences
CC	(AAU02435-AAU02452) are described in the present invention. The invention
CC	relates to a novel method of cloning cDNA encoding cell surface antigens
CC	and efficient construction of cDNA libraries. Also described are 2
CC	expression vectors (AAS03171, AAS03174) which provide high level
CC	expression in eukaryotic host cells. A genetically engineered cDNA
CC	sequence encoding the CD28 amino acid extracellular domain sequence
CC	(amino acids 1-134 given in AAU02437) and/or comprising nucleotides
CC	100-759, 154-555 or 154-759 of the CD28 cDNA sequence (AAS03175) is also
CC	new. The purified genes and proteins are useful for immunodiagnostic and
CC	immunotherapeutic applications, such as in the diagnosis and treatment of
CC	immune-mediated diseases, infections or disorders in animals and humans.
CC	Such diseases include immune deficiency diseases, diseases of immediate
CC	type of hypersensitivity, asthma, hypersensitivity pneumonitis, systemic
CC	lupus erythematosus, rheumatoid arthritis, acute and chronic
CC	inflammation, platelet disorders, plasma and other cell neoplasms,
CC	parastitic diseases, multiple sclerosis, Guillain-Barre syndrome and
CC	tissue and organ transplant rejection. The sequences can also be used to
CC	identify, isolate and purify other antibodies and antigens.
XX	
SO	Sequence 405 AA:
Query Match	100.0%; Score 666; DB 22; Length 405;
Best Local Similarity	100.0%; Pred. No. 2.5e-63;
Matches 117; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 WYHYHSEKMMNQRRRCRDNDYTDLVAIQNAELIELEKTLPERSRYWGIRIGGIW 60
Db	52 WCYHYSEKPMNQRRRCRDNDYTDLVAIQNAELIELEKTLPISRYWIGIRI 111
OY	61 TWVGNKSTEEAEKNMGDEPNKKNKEDCEVIYIKRKMDAGKMDDACHRIKAALC 117
Db	112 TWGINKSTEEAEKNMGDEPNKKNKEDCEVIYIKRKMDAGKMDDACHRIKAALC 168
RESULT 40	
AAR98110	
ID	AAR98110 standard; Protein; 371 AA.
XX	
AC	AAR98110;
XX	
DT	01-NOV-1996 (first entry)
XX	
DE	Human lymphocyte cell surface glycoprotein (HuLHR) variant.
XX	
KM	Immunoglobulin; transmembrane receptor; adhesion; targetting;
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW	immunomodulator; cell adhesion; graft rejection; inflammation;
KW	metastasis.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Region Location/Qualifiers
FT	1..37 /label= Signal region.
FT	38..154 /label= Lectin domain.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:17 : Search time 61.04 Seconds
(without alignments)
96.360 Million cell updates/sec

Title: US-09-119-209-2_COPY_160_193
Perfect score: 215
Sequence: 1 CQPMSCSGHGECEIINNHNHCNDVGYGPGQL 34

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 17294929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	205	95.3	385	4	Q9UJ43 homo sapien
2	181	84.2	376	6	Q28629 oryctolagus
3	159	74.0	372	11	063762 rattus norv
4	131	60.9	609	6	09GLFO canis faml
5	127	59.1	484	6	09SLG2 ovis aries
6	127	59.1	649	6	Q28657 oryctolagus
7	125	58.1	482	6	Q28982 sus scrofa
8	125	58.1	616	4	095509 homo sapien
9	125	58.1	740	4	095507 homo sapien
10	125	58.1	740	4	095508 homo sapien
11	123	57.2	610	6	095IG1 equus caball
12	122	56.7	646	6	Q28097 sus scrofa
13	121	56.3	485	6	Q9SLG3 odocoileus
14	119	55.3	754	6	Q28290 canis faml
15	98	45.6	1551	5	Q9NGV4 drosophila
16	98	45.6	3396	5	Q9VM55 drosophila

17	95	44.2	710	5	Q93563	Q93563 caenorhabd
18	94	43.7	112	4	Q9H557	Q9H557 homo sapien
19	93	43.3	642	13	P79941	P79941 xenopus lae
20	93	43.3	723	4	Q9UJ43	Q9UJ43 homo sapien
21	93	43.3	723	4	Q9NU41	Q9NU41 homo sapien
22	91	42.3	2656	5	Q9GNUM3	Q9GNUM3 paracentrot
23	90	41.9	721	13	Q91902	Q91902 xenopus lae
24	90	41.9	2531	5	016004	016004 lytechinus
25	90	41.9	3767	5	Q9UA13	Q9UA13 caenorhabd
26	89.5	41.6	2524	5	Q9GPA5	Q9GPA5 branchiosto
27	89	41.4	1203	11	Q06008	Q06008 mus musculo
28	89	41.4	2281	4	Q9UPL3	Q9UPL3 homo sapien
29	89	41.4	2319	11	Q9UPL7	Q9UPL7 rattus norv
30	89	41.4	2321	4	Q9Y6L8	Q9Y6L8 homo sapien
31	89	41.4	2321	4	Q9UM47	Q9UM47 homo sapien
32	89	41.4	2470	11	Q95516	Q95516 mus musculo
33	89	41.4	2471	4	Q04721	Q04721 homo sapien
34	89	41.4	2471	4	Q9H240	Q9H240 homo sapien
35	89	41.4	2471	11	Q9OW30	Q9OW30 rattus sp..
36	89	41.4	2653	5	Q25253	Q25253 lucilia cup
37	88.5	41.2	3209	13	Q93574	Q93574 gallus gall
38	88	40.9	529	5	Q25058	Q25058 helioicidari
39	88	40.9	562	5	Q9VQ12	Q9VQ12 drosophila
40	88	40.9	585	5	Q9UOE2	Q9UOE2 tribolium c
41	88	40.9	721	5	Q95Y60	Q95Y60 clona savig
42	87	40.5	156	5	Q26661	Q26661 strongyloce
43	87	40.5	174	5	Q9NA32	Q9NA32 caenorhabd
44	87	40.5	460	11	Q99PC6	Q99PC6 mus musculo
45	87	40.5	460	11	Q91WN8	Q91WN8 mus musculo

ALIGNMENTS

RESULT 1
ID Q9UJ43 PRELIMINARY: PRT: 385 AA.
AC Q9UJ43:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE L-SELECTIN PRECURSOR.
GN L-SELECTIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);
RA Fieger C.B.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);
RA Fieger C.B.;
RL Thesis (1998), Freie Universitaet Berlin, Fachbereich Chemie.
DR EMBL: AJ246000; CAB5488.1; -.
DR HSSP: P1451; IKUB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; Lectin.C.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sush1_2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP_2.
DR SMART: SM00034; CUECT_1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE: PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00181; EGF; 4.
DR SMART; SM00001; EGF_1like; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW Lectin; Selectin.
SQ SEQUENCE 609 AA; 66073 MW; 41E62D1F4D23881F CRC64;

Query Match 60.9%; Score 131; DB 6; Length 609;
Best Local Similarity 60.6%; Pred. No. 5,4e-10;
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 1 COPMSCSGHCEVEITNNHTCNCVDGYGPOCQ 33
DB 144 CPTSCSGHCEVEITNNHTCNCCHPFGKLRCE 176

RESULT 5
OY5LG2 PRELIMINARY; PRT; 484 AA.

AC 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE E-SELECTIN.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21421234; PubMed-11529941;
RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,
RA MacLachlan N.J.;
RT "Characterization of equine E-selectin."
RL Immunology 103:498-504(2001).
DR EMBL: AF307971; AAK4871.1; -.
KW Lectin; Selectin.
SQ SEQUENCE 484 AA; 52951 MW; 617930C1C2F47B44 CRC64;

Query Match 59.1%; Score 127; DB 6; Length 484;
Best Local Similarity 60.6%; Pred. No. 1,5e-09;
Matches 20; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 COPMSCSGHCEVEITNNHTCNCVDGYGPOCQ 33
DB 143 CPTSCSGHCEVEITNNHTCNCCHPFGKLRCE 175

RESULT 6
OY8657 PRELIMINARY; PRT; 649 AA.
AC 028657;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P-SELECTIN.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Vora D.K., Fang Z., Liva S.M., Parham F., Watson A.D., Drake T.A.,
RA Terfio M.C., Berlinger J.A.;
RT "Induction of P-selectin by MM-IDL and its role in human
RT atherosclerosis."
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Warden C.H.;

RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39446; AAA81385.1; -.
DR HSSP; P16109; 1FSB.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00059; Lectin_C; 1.
DR Pfam; PF00084; Sush1; 6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 6.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein; Lectin; Selectin.
SQ SEQUENCE 649 AA; 71755 MW; ECCD8C847B84BC31 CRC64;

Query Match 59.1%; Score 127; DB 6; Length 649;
Best Local Similarity 57.6%; Pred. No. 2,1e-09;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 COPMSCSGHCEVEITNNHTCNCVDGYGPOCQ 33
DB 163 CODMSCSKQGETETIGNYTCSCYPGFSPECE 195

RESULT 7
OY8982 PRELIMINARY; PRT; 482 AA.

AC 028982;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE E-SELECTIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97075911; PubMed-8918234;
RA Winkler H., Brostjan C., Csizmadia V., Natarajan G., Anrather J.,
RA Bach F.H.;
RT "The Intron-exon structure of the porcine E-selectin-encoding gene."
RL Gene 176:67-72(1996).
DR EMBL; U37521; AAC48680.1; -.
DR HSSP; P16581; 1ESL.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00059; Lectin_C; 1.
DR Pfam; PF00084; Sush1; 4.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 4.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein; Lectin; Selectin.
SQ SEQUENCE 482 AA; 52341 MW; 97DC5D70BF115944 CRC64;

Query Match 58.1%; Score 125; DB 6; Length 482;
Best Local Similarity 60.6%; Pred. No. 2,9e-09;

	Matches	20, Conservative	4, Mismatches	9, Indels	0, Gaps
QY	1	CQPMSCSGHGECVELIINNHNTCMCDVGIYQPCQ	33		
Db	144	CTPTSCSGHGECIELTINSTCCQYGFGRGLQCE	176		

RESULT	ID	PRELIMINARY	PRT	616 AA.
095509	095509	PRELIMINARY	PRT	616 AA.
AC	095509			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)			
DE	DJ780M13.1.3 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN			
DE	CD62, GMP140)) (ISOFORM 3) (FRAGMENT).			
GN	SELP.			
OC	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Howden P.;			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
RL	EMBL, AL022146; CAA18144.1; -			
DR	HSSP; P16109; 1KTD.			
DR	Interpro: IPR000561; EGF-like.			
DR	Interpro: IPR001304; lectin_c.			
DR	Interpro: IPR002396; Selectin.			
DR	Interpro: IPR000436; Sushi_SCR_CCP.			
DR	Pfam; PF00008; EGF_1.			
DR	Pfam; PF00059; lectin_c_1.			
DR	Pfam; PF00084; sushi_6.			
DR	PRINTS; PR00343; SELECTIN.			
DR	SMART; SM00037; CCP_6.			
DR	SMART; SM00034; CLECT_1.			
DR	SMART; SM00181; EGF_1.			
DR	PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.			
DR	PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.			
DR	PROSITE; PS01186; EGF_2; 1.			
KW	EGF-like domain; Glycoprotein.			
FT	NON_TER			
FT	616			
FT	616			
SEQ	SEQUENCE	616 AA;	67736 MW;	35CD4BFAD61D724 CRC64;
Query Match		58.1%;	Score 125;	DB 4; Length 616;
Best Local Similarity		57.6%;	Pred. No. 3.7e-09;	
Matches	19; Conservative	6;	Mismatches	8; Indels 0; Gaps
OY	1	COPMSCGHEGCEVETIINHNTGNCDDVGYGPOCQ	33	
Db	163	CODMSCSKGEGCLFTIGNYTSCYCPGFPECE	195	
RESULT	9			
095507	095507	PRELIMINARY	PRT	740 AA.
AC	095507			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)			
DE	DJ780M13.1.1 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN			
DE	CD62, GMP140)) (ISOFORM 1) (FRAGMENT).			
GN	SELP.			
OC	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Howden P.;			

RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases
 DR EMBL: AL022146; CAA181442.1; -
 DR HSSP; P16109; 1KTD
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001304; Ilectin_c.
 DR InterPro; IPR002396; Selectin.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00059; Ilectin_c.
 DR Pfam; PF00084; sushi_8.
 DR PRINTS; PR00343; SRILECTIN.
 DR SMART; SM00032; CCP; 8.
 DR SMART; SM00034; CLEC7_1.
 DR SMART; SM00181; EGF_1.
 DR PROSITE; PS00615; C_Type_Lectin_1; 1.
 DR PROSITE; PS50041; C_Type_Lectin_2; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 1
 KW EGF-like domain; glycoprotein.
 FT
 NON_TER 740 740
 SQ SEQUENCE 740 AA; 81095 MW; 3B5F70A45B1A3CD4 CRC64;

Query Match	58.1%;	Score 125;	DB 4;	Length 740;
Best Local Similarity	57.6%;	Pred. No. 4.5e-09;		
Matches 19;	Conservative 6;	Mismatches 8;	Indels 0;	Gaps 0;

OY 1 CQPMSCSGHGECEYEIINNHTCNCDDVGYGPOCQ 33
 || ||| ||| : | : || : | | : ||| :
Db 163 QDMSCKQGECLETTIGNYTCSCTPFGYGPCE 195

RESULT	10			
095508				
ID	095508	PRELIMINARY;	PRT;	740 AA.
AC	095508;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	DJ780M13.1.2 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN			
DE	362, GMP140)) (ISOFORM 2) (FRAGMENT).			
GN	SELP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxId=9606;			
KN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Howden P.;			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL022146; CAA18143.1; -.			
DR	HSSP; P16109; 1KTD.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001304; lectin_c.			
DR	InterPro; IPR002396; Selectin.			
DR	InterPro; IPR00436; Sush1_SCR_CCP.			
DR	Pfam; PF00008; EGF; 1.			
DR	Pfam; PF00059; lectin_c; 1.			
DR	Pfam; PF00084; sush1; 8.			
DR	PRINTS; PR00343; SELECTIN.			
DR	SMART; SM00032; CCP_8.			
DR	SMART; SM00034; CLECT; 1.			
DR	SMART; SM00181; EGF; 1.			
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.			
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.			
DR	PROSITE; PS01022; EGF_1; UNKNOWN_1.			
DR	PROSITE; PS01186; EGF_2; 1.			
KW	EGF-like domain; Glycoprotein.			
FT	NON_TER	740	740	
Q	SEQUENCE	740 AA;	81390 MW;	1D2E35BD93745CE CRC64;

Query Match 58.18; Score 125; DB 4; Length 740;

Best Local Similarity 57.6%; Pred. No. 4,5e-09;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33
DB 163 CQSTSCSKGECRETIGNYTCSCYSGFYGPCE 195

RESULT 11

OY 095LG1 PRELIMINARY; PRT; 610 AA.
AC 095LG1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE E-SELECTIN.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21421234; PubMed=11529941;
RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,
RA MacLachlan N.J.;
RT "Characterization of equine E-selectin."
RL Immunology 103:498-504(2001).
DR EMBL; AF307972; AAK48712.1; -.
KW Lectin; Selectin.
SQ SEQUENCE 610 AA; 66191 MW; F9D3DED12C445382 CRC64;

Query Match 57.2%; Score 123; DB 6; Length 610;

Best Local Similarity 60.6%; Pred. No. 6,9e-09;
Matches 20; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33
DB 142 CHTSCSGHGECEIINNHTCNCNDVGYGPOCQ 174

RESULT 12

OY 029097 PRELIMINARY; PRT; 646 AA.
AC 029097;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE P-SELECTIN PRECURSOR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=AORTA;
RA Rollins S.A., Johnson K.K., Birks C.W., Matlis L.A., Rother R.P.;
RN Submitted (Oct-1995) to the EMBL/Genbank/DBD databases.
RL [2]
RP SEQUENCE FROM N.A.
RC TISSUE=AORTA;
RX MEDLINE=20171534; PubMed=10706724;
RA Stocker C.J., Sugars K.L., Harari O.A., Landis R.C., Morley B.J.,
RA Haskard D.O.;
RT "TNF-alpha, IL-4, and IFN-gamma regulate differential expression of P-
RT and E-selectin expression by porcine aortic endothelial cells."
RL J. Immunol. 164:3309-3315(2000).
DR EMBL; L39075; AAA79007.1; -.
DR EMBL; AF163766; AAF43272.1; -.
DR HSP; P16109; IFSB.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001304; lectin_C.
DR InterPro; IPR002396; Selectin.

DR InterPro; IPR000436; Sush1_SCR_CCP.

DR Pfam; PF00008; EGF_1.
DR Pfam; PF00059; lectin_C_1.
DR Pfam; PF00084; sush1_6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00034; CCP; 6.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.
FT SIGNAL 1 41 POTENTIAL.
FT CHAIN 42 646 P-SELECTIN.
SQ SEQUENCE 646 AA; 71127 MW; 3863FAFE09F0BB6 CRC64;

Query Match 56.7%; Score 122; DB 6; Length 646;

Best Local Similarity 57.6%; Pred. No. 1e-08;
Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33
DB 163 CQSTSCSKGECRETIGNYTCSCYSGFYGPCE 195

RESULT 13

OY 095LG3 PRELIMINARY; PRT; 485 AA.
AC 095LG3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE E-SELECTIN.
OS Odocoileus hemionus (Mule deer) (Black-tailed deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Odocoileinae; Odocoileus.
OX NCBI_TaxID=9872;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21421234; PubMed=11529941;
RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,
RA MacLachlan N.J.;
RT "Characterization of equine E-selectin."
RL Immunology 103:498-504(2001).
DR EMBL; AF307970; AAK48710.1; -.
KW Lectin; Selectin.
SQ SEQUENCE 485 AA; 53247 MW; 69959199EAFE9980 CRC64;

Query Match 56.3%; Score 121; DB 6; Length 485;

Best Local Similarity 57.6%; Pred. No. 1e-08;
Matches 19; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33
DB 144 CHTPCSSHGECEIINNHTCNCNDVGYGPOCQ 176

RESULT 14

OY 028290 PRELIMINARY; PRT; 754 AA.
AC 028290;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CELL ADHESION MOLECULE PRECURSOR (FRAGMENT).
GN GMP140.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OK NCBI_TaxID=9615;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Manning A.M., Sanders W.E.Jr., Kukielka G.L., Dore M.,
RA Rosenbloom C.L., Hawkins H.L., Michael L.H., Ertman M.L., Smith C.W.,
RA Beaudet A.L., Anderson D.C.,
RT "Molecular cloning of canine GMP140 and studies of expression in a
RT model of myocardial ischemia/reperfusion."
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: M88170; AAA63789.1; -
DR HSSP: P16109; 1FSB
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR002396; Selectin.
DR Pfam: PF00059; lectin_c.1.
DR Pfam: PF00084; sushi; 8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
DR EGF-like domain; Glycoprotein; Signal.
FT NON_TER 1
FT SIGNAL <1 13 POTENTIAL.
FT CHAIN 14 754 CELL ADHESION MOLECULE.
FT SEQUENCE 754 AA; 82303 MW; F0438BEAA521E773 CRC64;

Query Match 55.3%; Score 119; DB 6; Length 754;
Best Local Similarity 54.5%; Pred. No. 3.1e-08;
Matches 18; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
OY 1 CQWSCSGHGCVEIINHTQNCVGYGPQCQ 33
DB 148 CQWSCSKQGECLFTIGWTCSPFGYSCCE 180

RESULT 15
O9NGV4 PRELIMINARY; PRT; 1551 AA.
AC O9NGV4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SP1070.
GN SP1070 OR CG9138.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENDOPLASMIC RETICULUM;
RA Serrano T.L., Pendleton J.D., Rubin G.M.;
RT "A reverse genetic screen for genes involved in Drosophila
RT development."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF239608; AAF63500.1; -
DR HSSP: P00740; 1EDM.
DR FlyBase; Fggn0031879; SP1070.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF-IT.
DR InterPro: IPR003410; HYR.
DR InterPro: IPR002049; Laminin_EGF.

DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR002172; LDL_recept_A.
DR Pfam: PF00008; EGF; 16.
DR Pfam: PF02494; HYR; 1.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00011; EGFBLAMININ.
DR SMART: SM00179; EGF_CA; 6.
DR SMART: SM00001; EGF-like; 9.
DR SMART: SM00282; LamC; 1.
DR SMART: SM00182; LDLa; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_9.
DR PROSITE: PS00022; EGF_1; UNKNOWN_15.
DR PROSITE: PS01186; EGF_2; 12.
DR PROSITE: PS01187; EGF_CA; 5.
KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
SO SEQUENCE 1551 AA; 167816 MW; A97EA229E9384F31 CRC64;

Query Match 45.6%; Score 98; DB 5; Length 1551;
Best Local Similarity 39.4%; Pred. No. 5.3e-05;
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
OY 1 CQWSCSGHGCVEIINHTQNCVGYGPQCQ 33
DB 251 CENRPCANEGTCVDLVGYSCNCEPGYTKNQC 283

RESULT 16
O9VM55 PRELIMINARY; PRT; 3396 AA.
AC O9VM55;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG9138 PROTEIN.
GN SP1070 OR CG9138.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Durbin K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1 SIMILARITY: CONTAINS 3 CUB DOMAINS.
DR EMBL: AE003615; AAF52472.1; -.
DR HSSP: P00740; 1EDM.
DR FlyBase: FBgn0031879; SP1070.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000859; CUB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_IT.
DR InterPro: IPR000421; PA58_C.
DR InterPro: IPR001092; HLH_dim.
DR InterPro: IPR003410; HVR.
DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR000436; Sushl_SCR_CCP.
DR InterPro: IPR001368; TNFR_C6.
DR Pfam: PF00431; CUB; 3.
DR Pfam: PF00008; EGF; 16.
DR Pfam: PF00754; F5_F8_type_C; 2.
DR Pfam: PF02494; HVR; 3.
DR Pfam: PF00057; Idl_recept_a; 1.
DR Pfam: PF00084; sushl; 7.
DR PRINTS: PR00010; EGFBLDOD.
DR SMART: SM0032; CCP; 8.
DR SMART: SM0042; CUB; 3.
DR SMART: SM00179; EGF_CA; 8.
DR SMART: SM00001; EGF-like; 9.
DR SMART: SM00231; PA58C; 2.
DR SMART: SM00282; LamG; 1.
DR SMART: SM00192; LDla; 1.
DR SMART: SM00208; TNFR; 2.
DR PROSITE: PS00010; ASX_HKROXYL; 11.
DR PROSITE: PS01180; ASX; 6.
DR PROSITE: PS00022; EGF_1; 15.
DR PROSITE: PS01186; EGF_2; 13.
DR PROSITE: PS01187; EGF_CA; 7.
DR PROSITE: PS01285; PA58C_1; 1.
DR PROSITE: PS01209; LDLa_1; 1.
DR PROSITE: PS00068; LDLa_2; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 3396 AA; 369389 MW; E618E9ACEA13E0E5 CRC64;

Query Match 45.6%; Score 98; DB 5; Length 3396;
Best Local Similarity 39.4%; Pred. No. 0.00012;
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECVELINHTNCNDVGYGPOCQ 33
Db 2092 CENRPCANEGTCVLDVGYSCNCEPGYTGKNQ 2124

RESULT 17
Q93563 PRELIMINARY; PRT; 710 AA.
AC 093563;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F25D7.5 PROTEIN.
GN F25D7.5.
OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodertinae; *Caenorhabditis*.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurry A.A.;
RL Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode *C. elegans*: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: 278418; CAB01698.1; -.
DR InterPro: IPR000561; EGF-like.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 710 AA; 80783 MW; 4F5DF8112C5D558 CRC64;

Query Match 44.2%; Score 95; DB 5; Length 710;
Best Local Similarity 38.2%; Pred. No. 6.2e-05;
Matches 13; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECVELINHTNCNDVGYGPOCQ 34
Db 148 CPTTCNGHKCYDVEDKDCQWYEGEHCEV 181

RESULT 18
Q9H557 PRELIMINARY; PRT; 112 AA.
AC Q9H557;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DJ303F19.1 (SIMILAR TO NOTCH (DROSOPHILA) HOMOLOG) (FRAGMENT).
GN DJ303F19.1.
OS *Homo sapiens* (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams S.;
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL137007; CAC10575.1; -.
DR HSSP: P00749; TURK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_IT.
DR PRINTS: PR00010; EGFBLDOD.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM00001; EGF-like; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12638 MW; 4141BD74D9924E7A CRC64;

Query Match 43.7%; Score 94; DB 4; Length 112;
Best Local Similarity 41.2%; Pred. No. 1.3e-05;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECVELINHTNCNDVGYGPOCQ 34
Db 47 CSSEPCKNNGTCVLDLTNRFNCCEPEYHGPFCE 80

```
RESULT 19
P79941 PRELIMINARY; PRT: 642 AA.
ID AC P79941;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NOTCH LIGAND X-DELTA-2.
GN X-DELTA-2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBL_TaxID=6355;
RN [1]
RP SEQUENCE FROM N.A.
RA Jen W.C., Weltstein D.A., Chitnis A.B., Kintner C.;
RT "The Notch ligand, X-Delta-2, mediates segmentation of the paraxial
RT mesoderm in Xenopus embryos."
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR HSSP: P00740; 1EDM.
DR PROSITE: PS00015; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_TI.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PF00008; EGF; 8.
DR PRINTS: PR00010; EGFBL00D.
DR SMART: SM00051; DSL; 1.
DR SMART: SM00179; EGF_CA; 3.
DR SMART: SM00001; EGF_like; 5.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 7.
DR PROSITE: PS01187; EGF_CA; 2.
DR Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
KW SEQUENCE 642 AA; 70667 MW; D7DC31EB9D92820C CRC64;
SO

Query Match 43.3%; Score 93; DB 13; Length 642;
Best Local Similarity 41.2%; Pred. No. 0.00011;
Matches 14; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

OY 1 CQWSSGSGHCEYEIINHTNCNDGVYGPQCQL 34
DB 428 CASSPCANGSTCVDAVNSTYSCSTGLAYGGKCTL 461

RESULT 20
O9UVJ2 PRELIMINARY; PRT: 723 AA.
ID AC O9UVJ2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE DELTA-LIKE-1 PROTEIN.
GN DLI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Han W., Ye Q., Moore M.A.S.;
RT "A Soluble Form of Human Delta-Like-1 Inhibits Differentiation of
RT Hematopoietic Progenitor Cells."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF196571; AAF05834.1; -.
```

```
DR HSSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_TI.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PF00008; EGF; 8.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00051; DSL; 1.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_like; 4.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 7.
DR PROSITE: PS01187; EGF_CA; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
KW SEQUENCE 723 AA; 77985 MW; C7F41BDA2265A5F7 CRC64;
SO
```

Query Match 43.3%; Score 93; DB 4; Length 723;
Best Local Similarity 38.2%; Pred. No. 0.00012;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

```
OY 1 CQWSSGSGHCEYEIINHTNCNDGVYGPQCQL 34
DB 332 CDSPCKNGSCTDLNYSCTCPGFGKICEL 365

RESULT 21
O9NU41 PRELIMINARY; PRT: 723 AA.
ID AC O9NU41;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE DJ894D12.3 (DELTA-LIKE 1 (MOUSE) HOMOLOG).
GN DLI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Almeida J.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL078605; CAB89569.1; -.
DR HSSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_TI.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PF00008; EGF; 8.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00051; DSL; 1.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_like; 4.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 8.
DR PROSITE: PS01187; EGF_CA; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
KW SEQUENCE 723 AA; 78055 MW; 0948BF25DDP899D CRC64;
SO
```

```

Query Match          43.3%; Score 93; DB 4; Length 723;
Best Local Similarity 38.2%; Pred. No. 0.00012;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINHTCNCMDVGYGPQCQL 34
DB 332 CDSPCKNGSCSDLNSYTCSPGFGKICEL 365

RESULT 22
Q9GNUM3 PRELIMINARY: PRT; 2656 AA.
ID Q9GNUM3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE FIBROSURFIN PRECURSOR.
GN SURFIN2656.
OS Paracentrotus lividus (Common sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinidae;
OC Paracentrotus.
OX NCBI_TaxID=7656;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADULT TEST;
RX MEDLINE=21264600; PubMed=11259425;
RA Clavel C., Lechias C., Humbert F., Garrone R., Exposito J.Y.;
RT "Characterization of fibrosurin, an interfibrillar component of sea
RT urchin catch connective tissues.";
RL J. Biol. Chem. 276:18108-18114(2001).
DR EMBL: AJ291489; CAC20782.1; -.
DR HSSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR001438; EGF-II.
DR Pfam: PF00008; EGF; 17.
DR PRINTS: PR00010; EGFBLD.
DR SMART: SM00181; EGF; 17.
DR SMART: SM00179; EGF-Ca; 17.
DR SMART: SM00001; EGF-like; 5.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
SO SEQUENCE 2656 AA; 276689 MW; 7CEC6A190886C229 CRC64;

Query Match          42.3%; Score 91; DB 5; Length 2656;
Best Local Similarity 39.4%; Pred. No. 0.00087;
Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINHTCNCMDVGYGPQCQL 33
DB 2556 CASGPCENAGDCIDDEVNSTCCTAGYBELVQ 2588

RESULT 23
Q91902 PRELIMINARY: PRT; 721 AA.
ID Q91902;
AC Q91902;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE X-DELTA-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=95319507; PubMed=7596411;
RA Henrique D., Adam J., Wrat A., Chitnis A., Lewis J., Ish-Horowitz D.;
RT "Expression of a Delta homologue in prospective neurons in the
RT chick.";
RL Nature 375:787-790(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95319503; PubMed=7596407;
RA Chitnis A., Henrique D., Lewis J., Ish-Horowitz D., Kintner C.;
RT "Primary neurogenesis in Xenopus embryos regulated by a homologue of
RT the Drosophila neurogenic gene Delta.";
RL Nature 375:761-766(1995).
DR EMBL: L42229; AAC38017.1; -.
DR HSSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR001438; EGF-II.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PF00008; EGF; 8.
DR PRINTS: PR00010; EGFBLD.
DR SMART: SM00051; DSL; 1.
DR SMART: SM00179; EGF-Ca; 4.
DR SMART: SM00001; EGF-like; 4.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS00022; EGF-1; UNKNOWN_8.
DR PROSITE: PS01186; EGF-2; 8.
DR PROSITE: PS01187; EGF-Ca; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
SO SEQUENCE 721 AA; 79922 MW; 9EBDC85C439DD194 CRC64;

Query Match          41.9%; Score 90; DB 13; Length 721;
Best Local Similarity 38.2%; Pred. No. 0.00031;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINHTCNCMDVGYGPQCQL 34
DB 334 CDANPCKNGSCSDLNSYTCSPGFGKNCLEL 367

RESULT 24
O16004 PRELIMINARY: PRT; 2531 AA.
ID O16004;
AC O16004;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH HOMOLOG.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7654;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97454256; PubMed=9310331;
RA Sherwood D.R., McClay D.R.;
RT "Identification and localization of a sea urchin Notch homologue:
RT insights into vegetal plate regionalization and Notch receptor
RT development.";
RL Development 124:3363-3374(1997).
DR EMBL: AF000634; AAB82088.1; -.
DR HSSP: P01132; 1EGF.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR001438; EGF-II.
DR InterPro: IPR000800; Notch.

```

DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00008; EGF; 35.
 DR Pfam: PF00066; notch; 3.
 DR PRINTS: PR01415; ANKYRIN.
 DR PRINTS: PR00010; EGFBLD.
 DR PRINTS: PR01452; NOTCH.
 DR SMART: SM00248; ANK; 5.
 DR SMART: SM00179; EGF_CA; 23.
 DR SMART: SM00004; EGF_like; 11.
 DR SMART: SM00004; NL; 3.
 DR PROSITE: PS50088; ANK_REPEAT; 10.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 21.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_33.
 DR PROSITE: PS01186; EGF_2; 25.
 DR PROSITE: PS01187; EGF_CA; 20.
 DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;
 KW Hydroxylation; Repeat.
 SO SEQUENCE 2531 AA; 273982 MW; 5BF42BEC627CA303 CRC64;

Query Match 41.9%; Score 90; DB 5; Length 2531;
 Best Local Similarity 38.2%; Pred. No. 0.0011;
 Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 OY 1 CQPMSCGHECVCETINNNHTCNCDVGYGPOCOL 34
 Db 555 CQSPSCENGCTCIDGVNQTCTCEGTGYGHRCEM 588

RESULT 25
 ID Q9UAI3 PRELIMINARY; PRT; 3767 AA.
 AC Q9UAI3; Q21340;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE TRANSMEMBRANE CELL ADHESION RECEPTOR MUA-3 PRECURSOR (K08E5.3 PROTEIN)
 DE (FRAGMENT).
 GN MUA-3 OR K08E5.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NZ;
 RA Lu Z., Vogel B., Hedgecock E.;
 RT "mua-3 mRNA Splicing Pattern Revealed";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kershaw J.K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: AF139060; AMD29428.1; -
 DR EMBL: Z30974; CAA83226.2; -
 DR EMBL: Z30423; CAA83226.2; JOINED.
 DR HSSP; P01130; ILDL.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001762; Disintegrin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR002172; LDL_recept_A.
 DR InterPro: IPR000082; SEA.
 DR InterPro: IPR002035; VWFA.
 DR Pfam: PF00008; EGF; 33.

DR Pfam: PF00057; Idl_recept_a; 4.
 DR Pfam: PF01390; SEA; 2.
 DR Pfam: PF00092; vwa; 1.
 DR PRINTS: PR00289; DISINTEGRIN.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR PRINTS: PR00453; VWFADOMAIN.
 DR SMART: SM00181; EGF; 51.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00001; EGF_like; 45.
 DR SMART: SM00192; LDLra; 4.
 DR SMART: SM00200; SEA; 2.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_32.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 6.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS50068; LDLRA_2; 2.
 DR PROSITE: PS50024; SEA; 4.
 DR PROSITE: PS50234; VWFA; 2.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Receptor; Repeat;
 KW Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CRAIN 25 >3767 TRANSMEMBRANE CELL ADHESION RECEPTOR MUA-
 FT 3.
 SO SEQUENCE 3767 AA; 417284 MW; 8DA3AE5EA50AE88E CRC64;

Query Match 41.9%; Score 90; DB 5; Length 3767;
 Best Local Similarity 51.6%; Pred. No. 0.0017;
 Matches 16; Conservative 4; Mismatches 9; Indels 2; Gaps 1;
 OY 6 CQSGHECVCETINNNHTCNCDVGYGPOCOL 34
 Db 3380 CNAHGDCVHTATNNTTCCTDGTGPOCOV 3410

RESULT 26
 ID Q9GPA5 PRELIMINARY; PRT; 2524 AA.
 AC Q9GPA5;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE NOTCH RECEPTOR PROTEIN.
 GN NOTCH.
 OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_TaxID=7739;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LARVA;
 RA Holland L.Z., Burgdorf C., Holland N.D., Lehrach H., Tamme R.,
 RA Abi-Rached L., Pontarotti P., Lardelli M.;
 RT "Cloning and developmental expression of the amphioxus homologue of
 RT Notch (Amphinotch): evolutionary conservation of multiple expression
 RT domains in amphioxus and vertebrates.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y12539; CAC19873.1; -
 DR HSSP; P00740; IEDM.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001438; EGF_11.
 DR InterPro: IPR002049; Laminin_EGF.
 DR InterPro: IPR000800; Notch.
 DR Pfam: PF00008; EGF; 36.
 DR Pfam: PF00066; notch; 3.
 DR PRINTS: PR00010; EGFBLD.
 DR PRINTS: PR00011; EGFLAMININ.


```
DR PRINTS: PRO1452; NOTCH.
DR SMART: SM00248; ANK: 6.
DR SMART: SM00181; EGF: 37.
DR SMART: SM00179; EGF_CA: 34.
DR SMART: SM00001; EGF_Like: 13.
DR SMART: SM00004; NL: 3.
DR PROSITE: PS50088; ANK_REPEAT: 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION: 1.
DR PROSITE: PS00010; ASX_HYDROXYL: 1.
DR ANK repeat: EGF-like domain; Glycoprotein; Hydroxylation; Receptor;
KW Repeat.
SQ SEQUENCE 2524 AA; 270969 MW; C2CA57E306D23EC9 CRC64;

Query Match 41.6%; Score 89.5; DB 5; Length 2524;
Best Local Similarity 45.5%; Pred. No. 0.0013;
Matches 15; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Oy 1 COPMSCSGHECEVIINHTCNCQVGYGPOCQ 33
Db 559 COSNCCQ-HGTCVGVASTCCEPGYNGPLCE 590

RESULT 27
006008 PRELIMINARY; PRT; 1203 AA.
AC 006008.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH B PROTEIN (FRAGMENT).
GN NOTCH2 OR NOTCH B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FL (CBA X C57BL); TISSUE=EMBRO;
RX MEDLINE=93178563; PubMed=8440332;
RA Lardelli M., Lendahl U.;
RT "Notch A and Notch B - two mouse Notch homologues coexpressed in a
RT wide variety of tissues."
RL Exp. Cell Res. 204:364-372(1993).
DR EMBL: X68279; CAA48340.1; -.
DR HSSP: P16109; IFSB.
DR MGD: MGI:97364; Notch2.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF_Like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_11.
DR InterPro: IPR000083; fibronectin_type_1.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00008; EGF_27.
DR Pfam: PF00066; notch; 1.
DR PRINTS: PRO0010; EGFLOOD.
DR PRINTS: PRO0011; EGF_LAMININ.
DR PRINTS: PRO0012; FNTYPEI.
DR SMART: SM00179; EGF_CA: 19.
DR SMART: SM00001; EGF_Like: 7.
DR SMART: SM00004; NL: 2.
DR PROSITE: PS00010; ASX_HYDROXYL: 19.
DR PROSITE: PS00022; EGF_1; UNKNOWN_27.
DR PROSITE: PS01186; EGF_2; 22.
DR PROSITE: PS01187; EGF_CA: 19.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW NON_TER 1
SQ SEQUENCE 1203 AA; 128962 MW; 65040843799BCA37 CRC64;

Query Match 41.4%; Score 89; DB 11; Length 1203;
```

```
Best Local Similarity 41.2%; Pred. No. 0.00072;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPMSCSGHECEVIINHTCNCQVGYGPOCQ 34
Db 181 COSNPCVNNQCVDKYNRFQCLCPGFTGPGVCQI 214

RESULT 28
09UPL3 PRELIMINARY; PRT; 2281 AA.
ID 09UPL3.
AC 09UPL3.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH 3 (FRAGMENT).
GN NOTCH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Ganes J., Danganan L., Poundstone P.,
RA Christensen M., Georgesou A., Avila J., Liu S., Atlix C., Andreise T.,
RA Bruce R., Thomas P., Quan G., Krommiller B., Arellano A.,
RA Montgomey M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S.,
RA Carrano A.V.;
RT "Sequence analysis of an ~1.5 Mb olfactory receptor (OLFR) cluster in
RT 19p13.1."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC004663; AAC15789.1; -.
DR HSSP: P00740; 1EDM.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR000561; EGF_Like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_11.
DR InterPro: IPR000083; fibronectin_type_1.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00008; EGF; 34.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PRO0289; DISINTEGRIN.
DR PRINTS: PRO0010; EGFLOOD.
DR PRINTS: PRO0011; EGF_LAMININ.
DR PRINTS: PRO0012; FNTYPEI.
DR SMART: SM00248; ANK: 5.
DR SMART: SM00179; EGF_CA: 19.
DR SMART: SM00001; EGF_Like: 14.
DR SMART: SM00004; NL: 3.
DR PROSITE: PS50088; ANK_REPEAT: 8.
DR PROSITE: PS50297; ANK_REPEAT_REGION: 2.
DR PROSITE: PS00010; ASX_HYDROXYL: 18.
DR PROSITE: PS00022; EGF_1; UNKNOWN_33.
DR PROSITE: PS01186; EGF_2; 25.
DR PROSITE: PS01187; EGF_CA: 16.
DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
KW NON_TER 1
SQ SEQUENCE 2281 AA; 239472 MW; DE23B2F129709D75 CRC64;

Query Match 41.4%; Score 89; DB 4; Length 2281;
Best Local Similarity 45.5%; Pred. No. 0.0014;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
```

```

Qy 1 CPMWCSGHGCEVEIINNHTCNDVGYGPQCQ 33
Db 888 CSPSSCFNGTCVGDVNSFSLCRPGYTGACQC 920

RESULT 29
Q9R172 PRELIMINARY; PRT; 2319 AA.
ID 09R172
AC 09R172;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NOTCH 3 PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haitunians T., Boulter J., Weimaster G., Schanen N.C.;
RT "Rattus norvegicus mRNA for Notch 3."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164486; AAD46653.2;
DR HSSP; P00740; 1EDM.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001438; EGF-II.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000800; Notch.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 33.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01452; NOTCH.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF-CA; 20.
DR SMART; SM00001; EGF-like; 12.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR ANK repeat; EGF-like domain; Glycoprotein; Repeat.
KW SEQUENCE 2319 AA; 244298 MW; 243BCA02D7C3283D CRC64;
SQ

Query Match 41.4%; Score 89; DB 11; Length 2319;
Best Local Similarity 45.5%; Pred. No. 0.0014;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

```

```

Qy 1 CPMWCSGHGCEVEIINNHTCNDVGYGPQCQ 33
Db 930 CSPSSCFNGTCVGDVNSFSLCRPGYTGACQC 962

RESULT 30
Q9Y6L8 PRELIMINARY; PRT; 2321 AA.
ID 09Y6L8
AC 09Y6L8;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NOTCH3.
GN NOTCH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Guel M., Artavanis-Tsakonas S.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF058900; AAC14346.1;
DR EMBL; AF058881; AAC14346.1; JOINED.
DR EMBL; AF058882; AAC14346.1; JOINED.
DR EMBL; AF058883; AAC14346.1; JOINED.
DR EMBL; AF058884; AAC14346.1; JOINED.
DR EMBL; AF058885; AAC14346.1; JOINED.
DR EMBL; AF058886; AAC14346.1; JOINED.
DR EMBL; AF058887; AAC14346.1; JOINED.
DR EMBL; AF058888; AAC14346.1; JOINED.
DR EMBL; AF058889; AAC14346.1; JOINED.
DR EMBL; AF058890; AAC14346.1; JOINED.
DR EMBL; AF058891; AAC14346.1; JOINED.
DR EMBL; AF058892; AAC14346.1; JOINED.
DR EMBL; AF058893; AAC14346.1; JOINED.
DR EMBL; AF058894; AAC14346.1; JOINED.
DR EMBL; AF058895; AAC14346.1; JOINED.
DR EMBL; AF058896; AAC14346.1; JOINED.
DR EMBL; AF058897; AAC14346.1; JOINED.
DR EMBL; AF058898; AAC14346.1; JOINED.
DR EMBL; AF058899; AAC14346.1; JOINED.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001438; EGF-II.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000800; Notch.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 34.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01452; NOTCH.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF-CA; 19.
DR SMART; SM00001; EGF-like; 14.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 18.
DR PROSITE; PS00022; EGF 1; UNKNOWN_33.
DR PROSITE; PS01186; EGF 2; 25.
DR PROSITE; PS01187; EGF-CA; 16.
KW ANK repeat; Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2321 AA; 243628 MW; 3E70EC12A59CD638 CRC64;

```

```

Query Match 41.4%; Score 89; DB 4; Length 2321;
Best Local Similarity 45.5%; Pred. No. 0.0014;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CPMWCSGHGCEVEIINNHTCNDVGYGPQCQ 33
Db 928 CSPSSCFNGTCVGDVNSFSLCRPGYTGACQC 960

RESULT 31
Q9UM47 PRELIMINARY; PRT; 2321 AA.
ID 09UM47
AC 09UM47;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NOTCH3.
GN NOTCH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97032728; PubMed=8678478;
RA Joutel A., Corpechot C., Ducros A., Vahedi K., Chabriat H., Mouton P.,
RA Almomtch S., Domenga V., Cecillon M., Marechal E., Maciazek J.,
RA Vaysiere C., Cruaud C., Cabanis E.A., Ruchoux M.M., Weissbach J.,
RA Bach J.F., Bousser M.G., Tournier-Lasserre E.;
RT "Notch3 mutations in CADASIL, a hereditary adult-onset condition
RT causing stroke and dementia.";
RL Nature 383:707-710(1996).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98049753; PubMed=9388399;
RA Joutel A., Vahedi K., Corpechot C., Trosch A., Chabriat H.,
RA Vaysiere C., Cruaud C., Maciazek J., Weissbach J., Bousser M.G.,
RA Bach J.F., Tournier-Lasserre E.;
RT "Strong clustering and stereotyped nature of Notch3 mutations in
RT CADASIL patients.";
RL Lancet 350:1511-1515(1997).
DR EMBL: U97669; AAB91371.1; -.
DR HSSP: P00740; IEDM.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00008; EGF; 34.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00010; EGFBLDOD.
DR PRINTS: PR00011; EGFBLAMIN.
DR PRINTS: PR01452; NOTCH.
DR SMART: SM00248; ANK; 5.
DR SMART: SM00179; EGF_CA; 19.
DR SMART: SM00001; EGF_Like; 14.
DR SMART: SM00004; NL; 3.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 18.
DR PROSITE: PS00022; EGF_1; UNKNOWN_33.
DR PROSITE: PS01186; EGF_2; 25.
DR PROSITE: PS01187; EGF_CA; 16.
KW ANK repeat; Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2321 AA; 243656 MW; AFABF2ECAB7C836 CRC64;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57B/6; TISSUE=THYMUS;
RX MEDLINE=93178563; PubMed=8440332;
RA Lardelli M., Lendahl U.;
RT "Notch A and Notch B - two mouse Notch homologues coexpressed in a
RT wide variety of tissues.";
RL Exp. Cell Res. 204:364-372(1993).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57B/6; TISSUE=THYMUS;
RA Hamada Y., Higuchi M., Tsujimoto Y.;
RT "Complete amino acid sequence and multiform transcripts encoded by a
RT single copy of mouse Notch2 gene.";
RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL: D32210; BAA22094.1; -.
DR HSSP: P16109; IFSB.
DR MGD: MGI:97364; Notch2.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00008; EGF; 35.
DR Pfam: PF00066; notch; 2.
DR PRINTS: PR00010; EGFBLDOD.
DR PRINTS: PR01452; NOTCH.
DR SMART: SM00248; ANK; 4.
DR SMART: SM00179; EGF_CA; 22.
DR SMART: SM00001; EGF_Like; 12.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 22.
DR PROSITE: PS00022; EGF_1; UNKNOWN_33.
DR PROSITE: PS01186; EGF_2; 27.
DR PROSITE: PS01187; EGF_CA; 22.
KW ANK repeat; Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2470 AA; 265325 MW; B5A31B35242716D CRC64;
```

Query Match 41.4%; Score 89; DB 11; Length 2470;
Best Local Similarity 41.2%; Pred. No. 0.0015;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

```
OY 1 CQWSCSGHGECEVEIINHTCMCDVGYGPQCL 34
Db 496 COSNPCVNNGCQVDKYNRFQCLCPGFTGPVCOI 529
[1]
ID 004721 PRELIMINARY; PRT; 2471 AA.
AC 004721;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE NOTCH 2 PRECURSOR.
GN N2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1810-2447 FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=93265135; PubMed=1303260;
RA Stifani S., Bismuth C.M., Redhead N.J., Hill R.E.,
RA Artavanis-Tsakonas S.;
```

```
RT "Human homologs of a Drosophila enhancer of split gene product define
RT a novel family of nuclear proteins."
RT Nat. Genet. 2:119-127(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RA Blumuelier C.M., Mann R.S.;
RT "Complete Human Notch 2 (hN2) cDNA sequence.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF308601; AAA36377.2; -.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00023; ank; 12.
DR Pfam; PF00008; EGF; 35.
DR Pfam; PF00066; notch; 2.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00248; ANK; 4.
DR SMART; SM00179; EGF_Ca; 22.
DR SMART; SM00004; NL; 2.
DR SMART; SM00001; EGF_like; 12.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR ANK repeat; EGF-like domain; Glycoprotein; Hydroxylation; Repeat;
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 2471 AA; 265401 MW; DB4FD003CCE5AF5 CRC64;

Query Match 41.4%; Score 89; DB 4; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.0015;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 CQWSCSGHGEVCEIINNHTCNCVDVGYGPQCL 34
Db 498 COSMPCVNGCVDKVNRFQCLCPGFTGPVCOI 531

RESULT 34
ID Q9H240 PRELIMINARY; PRT; 2471 AA.
AC Q9H240;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NOTCH2 PROTEIN.
GN NOTCH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST TUMOR;
RA Correa R.G., Camargo A.A., Moreira E.S., Simpson A.J.;
RT "Human Notch2, a novel member of cell-fate determining NOTCH family.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF315356; AAG37073.1; -.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00023; ank; 12.
DR Pfam; PF00008; EGF; 35.
DR Pfam; PF00066; notch; 2.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00248; ANK; 4.
DR SMART; SM00179; EGF_Ca; 35.
DR SMART; SM00001; EGF_like; 12.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR ANK repeat; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2471 AA; 265429 MW; 7DA7F8F44589B1C CRC64;
```

```
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000800; Notch.
DR Pfam; PF00008; EGF; 32.
DR Pfam; PF00066; notch; 2.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01452; NOTCH.
DR SMART; SM00248; ANK; 6.
DR SMART; SM00181; EGF; 37.
DR SMART; SM00179; EGF_Ca; 35.
DR SMART; SM00001; EGF_like; 12.
DR SMART; SM00004; NL; 2.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR ANK repeat; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2471 AA; 265429 MW; 7DA7F8F44589B1C CRC64;

Query Match 41.4%; Score 89; DB 4; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.0015;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 CQWSCSGHGEVCEIINNHTCNCVDVGYGPQCL 34
Db 498 COSMPCVNGCVDKVNRFQCLCPGFTGPVCOI 531

RESULT 35
ID Q9QW30 PRELIMINARY; PRT; 2471 AA.
AC Q9QW30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NOTCH2 PROTEIN.
GN Rattus sp., and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10118, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Rattus sp., and Rat; TISSUE=BRAIN;
RX MEDLINE=93202015; PubMed=1295745;
RA Weinmaster G., Roberts V.J., Lemke G.;
RT "Notch2: a second mammalian Notch gene.";
RL Development 116:931-941(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Rat; TISSUE=BRAIN;
RA Roberts V.J.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Rat; TISSUE=BRAIN;
RA Weinmaster G., Lemke G.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; M93661; AAK13558.1; -.
DR HSSP; P00743; 1CCF.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000800; Notch.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 35.
DR Pfam; PF00066; notch; 2.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGF_LAMININ.
```

PRINTS: PRO1452; NOTCH.
DR SMART: SM00248; ANK: 4.
DR SMART: SM00179; EGF_CA: 22.
DR SMART: SM00001; EGF_Like: 10.
DR SMART: SM00004; NL: 2.
DR PROSITE: PS50088; ANK_REPEAT: 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION: 1.
DR PROSITE: PS00010; ASX_HYDROXYL: 22.
DR PROSITE: PS00022; EGF_1; UNKNOWN_34.
DR PROSITE: PS01186; EGF_2; 26.
DR PROSITE: PS01187; EGF_CA: 22.
DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2471 AA; 265367 MW; 7D5C8E18DDE95F8 CRC64;

Query Match 41.4%; Score 89; DB 11; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.0015;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

OY 1 CPWCSGHCVEIINNHNTCNDVGYGPOCOL 34
DB 498 CQSNPCVNNGCVDKVNRFCCLPPTGTCVCOI 531

RESULT 36
025253 PRELIMINARY; PRT; 2653 AA.
AC 025253;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH HOMOLOG SCALLOPED WINGS (SCL).
GN SCL.
OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestridae; Calliphoridae; Lucilla.
OX NCBI_TaxID=7375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SS SEEKING;
RX MEDLINE=96400928; PubMed=8807304;
RA Davies A.G., Game A.Y., Chen Z., Williams T.J., Goodall S., Yen J.L.,
Mckenzie J.A., Batterham P.;
RT "Scalloped wings is the Lucilla cuprina Notch homologue and a
RT candidate for the modifier of fitness and asymmetry of diazino
RT resistance.";
RL Genetics 143:1321-1337(1996).
RN [2]
RP SEQUENCE OF 39-265 FROM N.A.
RC STRAIN=SS SEEKING;
RA Chen Z., Newsome T., McKenzie J.A., Batterham P.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 39-265 FROM N.A.
RC STRAIN=SS SEEKING;
RA Chen Z., McKenzie J.A., Batterham P.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U58977; AAC36151.1; -
DR EMBL: AF032672; AAC36152.1; -
DR EMBL: AF032671; AAC36152.1; JOINED.
DR EMBL: AF032671; AAC36152.1; JOINED.
DR EMBL: AF032673; AAC36153.1; -
DR HSSP: P00740; LEDM.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; ASX_HYDROXYL.
DR InterPro: IPR000561; EGF-Like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_11.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank; 6.

Pfam: PF00008; EGF: 36.
DR Pfam: PF00066; notch: 3.
DR PRINTS: PR00010; EGFLOD.
DR PRINTS: PRO1452; NOTCH.
DR SMART: SM00248; ANK: 4.
DR SMART: SM00179; EGF_CA: 24.
DR SMART: SM00001; EGF_Like: 11.
DR SMART: SM00004; NL: 2.
DR PROSITE: PS50088; ANK_REPEAT: 5.
DR PROSITE: PS50297; ANK_REPEAT_REGION: 1.
DR PROSITE: PS00010; ASX_HYDROXYL: 22.
DR PROSITE: PS00022; EGF_1; UNKNOWN_34.
DR PROSITE: PS01186; EGF_2; 28.
DR PROSITE: PS01187; EGF_CA: 21.
DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2653 AA; 285928 MW; 6AF2A058FEBC6329 CRC64;

Query Match 41.4%; Score 89; DB 5; Length 2653;
Best Local Similarity 45.2%; Pred. No. 0.0016;
Matches 14; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

OY 3 PWSGSGHCVEIINNHNTCNDVGYGPOCQ 33.
DB 1286 PCSNAGTLDVGLVNNHNCCKPGYMGRRCE 1316

RESULT 37
093574 PRELIMINARY; PRT; 3209 AA.
AC 093574;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EXTRACELLULAR REELIN (FRAGMENT).
GN RELN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Bernier B., Goffinet A.M.;
RT "Comparative study of reelin in vertebrates.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF090441; AAC35559.1; -
DR InterPro: IPR002860; BNR.
DR InterPro: IPR000561; EGF-Like.
DR Pfam: PF02012; BNR; 15.
DR Pfam: PF00008; EGF; 4.
DR SMART: SM00181; EGF; 5.
DR PROSITE: PS00022; EGF_1; UNKNOWN_7.
DR PROSITE: PS01186; EGF_2; 7.
KW EGF-like domain; Glycoprotein.
FT NON_TER 1
SQ SEQUENCE 3209 AA; 361288 MW; 81A7B6676BCA3D1 CRC64;

Query Match 41.2%; Score 88.5; DB 13; Length 3209;
Best Local Similarity 51.7%; Pred. No. 0.0023;
Matches 15; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

OY 6 CSGHGEVEIINNHNTCNDVGYGPOCOL 34
DB 1885 CMHGSC--INGTKICIDPGYSGPTCKI 1910

RESULT 38
025058 PRELIMINARY; PRT; 529 AA.
AC 025058;

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE FIBROPELIN IA (FRAGMENT).
OS Helicoidaris erythrogramma (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidea; Echinacea; Echinoidae; Echinometridae;
OC Helicoidaris.
OX NCBI_TaxID=7634;
RN [1]
RP SEQUENCE FROM N.A.
RA Bisgrove B.W.;
RT "Evolution of the fibropellin gene family and pattern of fibropellin
RT gene expression in sea urchin phylogeny."
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L33861; AAA2995.1; -.
DR HSP: P08709; 1BF9.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000088; Avidin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_T1.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF01382; Avidin_1.
DR PRINTS: PRO0010; EGFBLD.
DR PRINTS: PRO0011; EGFBLD.
DR PRINTS: PRO0012; FNTYPEI.
DR PRODOM: PD016055; Avidin; 1.
DR SMART: SM00179; EGF_CA; 9.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_8.
DR PROSITE: PS00577; AVIDIN; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_11.
DR PROSITE: PS01186; EGF_2; 10.
DR PROSITE: PS01187; EGF_CA; 7.
DR Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
FT NON_TER 1
SO SHODNCE 529 AA; 55543 MW; D4AE58FCF9ACB5A CRC64;

Query Match 40.9%; Score 88; DB 5; Length 529;
Best Local Similarity 36.4%; Pred. No. 0.0043;
Matches 12; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWCSGHEGVEIINHTCNCDCVGYGPOCQ 33
DB 215 CASMPLNGGCIEMWNGYTCOCAGFTGVLC 247

RESULT 39
Q9V012 PRELIMINARY; PRT; 562 AA.
AC Q9V012;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CG2991 PROTEIN.
GN CG2991.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S.P., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```

```

RA Brandon R.C., Rogers Y.-H.C., Blazew R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Asghar A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman J.T., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Modarity C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sytzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zibos X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003581; AAF51189.1; -.
DR FlyBase: FBgn0031474; CG2991.
DR InterPro: IPR000561; EGF-like.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
SO SEQUENCE 562 AA; 63026 MW; B377A00FA17761DA CRC64;

Query Match 40.9%; Score 88; DB 5; Length 562;
Best Local Similarity 50.0%; Pred. No. 0.0045;
Matches 15; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 5 SCSGHEGVEIINHTCNCDCVGYGPOCQ 34
DB 50 SCSGRDCV---NNTVCVDIRYAGNECDI 75

RESULT 40
Q9U0E2 PRELIMINARY; PRT; 585 AA.
AC Q9U0E2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIGNAL RECEPTOR PROTEIN (FRAGMENT).
GN NOTCH.
OS Tribolium castaneum (Red flour beetle).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Tenebrionidae; Tribolium.
OX NCBI_TaxID=7070;
RN [1]
RP SEQUENCE FROM N.A.
RA Tautz D., Lardelli M., Westlin J., Tamme R.;
RT "Embryonic expression of Tribolium Notch."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ005083; CAB65469.1; -.
DR HSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.

```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:41 : Search time 18.65 Seconds

(without alignments)
70.588 Million cell updates/sec

Title: US-09-119-209-2_COPY_160_193

Perfect score: 215

Sequence: 1 COPWSCSGHGEVCVEIINNHTCNCNDVGYRQCQL 34

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	205	95.3	372	1	LEM1_HUMAN
2	205	95.3	372	1	LEM1_HUMAN
3	205	95.3	372	1	LEM1_MACMU
4	205	95.3	372	1	LEM1_PANTR
5	205	95.3	372	1	LEM1_PAPHA
6	191	88.8	370	1	LEM1_PONPY
7	169	78.6	372	1	LEM1_BOVIN
8	159	74.0	372	1	LEM1_BOVIN
9	131	60.9	611	1	LEM1_RAT
10	130	60.5	610	1	LEM2_CANFA
11	127	59.1	549	1	LEM2_HUMAN
12	127	59.1	549	1	LEM2_RAT
13	127	59.1	646	1	LEM3_BOVIN
14	126	58.6	769	1	LEM3_SHEEP
15	125	58.1	612	1	LEM2_MOUSE
16	125	58.1	484	1	LEM2_PIG
17	124	57.7	768	1	LEM3_HUMAN
18	118	54.9	485	1	LEM3_MOUSE
19	116	54.0	551	1	LEM2_BOVIN
20	109	50.7	768	1	LEM3_RAT
21	95	44.2	603	1	FA12_CAVPO
22	93	43.3	723	1	DLL1_HUMAN
23	93	43.3	2524	1	NOTC_XENLA
24	90	41.9	3051	1	YMX3_CAEEL
25	87	40.5	461	1	PRTC_MOUSE
26	87	40.5	714	1	DLL1_RAT
27	87	40.5	1064	1	FBP1_STRPU
28	86	40.0	2139	1	CRB_DROME
29	86	40.0	2703	1	NOTC_DROME
30	85.5	39.8	473	1	FP2_MITGA
31	84	39.1	2201	1	TENA_HUMAN
32	84	39.1	2318	1	NTC3_MOUSE
33	83	38.6	461	1	PRTC_RAT

34	83	38.6	570	1	FBP3_STRPU
35	83	38.6	1376	1	CRBH_HUMAN
36	83	38.6	3097	1	CAON_DROME
37	82	38.1	2437	1	NOTC_BRARE
38	81	37.7	615	1	NOTC_HUMAN
39	81	37.7	799	1	ITBN_DROME
40	81	37.7	2444	1	NTC1_HUMAN
41	81	37.7	2531	1	NTC1_MOUSE
42	81	37.7	2531	1	NTC1_RAT
43	79	36.7	722	1	DLL1_MOUSE
44	79	36.7	1217	1	Egf_MOUSE
45	78.5	36.5	1964	1	NTC4_MOUSE

ALIGNMENTS

RESULT ID	1	LEM1_HUMAN	STANDARD:	PRT:	372 AA.
AC	P14151	P15023			
DT	01-JAN-1990	(Rel. 13, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte surface antigen Leu-8) (TQ1) (gp90-MEL)				
DE	(Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).				
GN	SEL1 OR LYAM1 OR LNHR.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89315837; PubMed=2664786;				
RA	Siegelman M.H., Weissman I.L.;				
RT	"Human homologue of mouse lymph node homing receptor: evolutionary conservation at tandem cell interaction domains.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:5562-5566(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=tonsil;				
RX	MEDLINE=89310350; PubMed=2473156;				
RA	Tedder T.F., Isaacs C.M., Ernst T.J., Demetri G.D., Adler D.A.,				
RT	Disteche C.M.;				
RT	"Isolation and chromosomal localization of cDNAs encoding a novel human lymphocyte cell surface molecule, LAM-1. Homology with the mouse lymphocyte homing receptor and other human adhesion proteins.";				
RL	J. Exp. Med. 170:123-133(1989).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90044046; PubMed=250939;				
RA	Camertini D., James S.P., Stamenkovic I., Seed B.;				
RT	"Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor.";				
RL	Nature 342:78-82(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89308881; PubMed=2663882;				
RA	Bower B.R., Nguyen T., Lasky L.A.;				
RT	"Characterization of a human homologue of the murine peripheral lymph node homing receptor.";				
RL	J. Cell Biol. 103:421-427(1989).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90243637; PubMed=1692315;				
RA	Ord D.C., Ernst T.J., Zhou L.J., Rambaldi A., Spertini O.,				
RT	Griffin J., Tedder T.F.;				
RT	"Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils.";				
RL	J. Biol. Chem. 265:7760-7767(1990).				
RN	[6]				
RP	3D-STRUCTURE MODELING.				

```

RX MEDLINE-96074584; PubMed-7488174;
RA Bajorath J., Anufro A.;
RT "A template for generation and comparison of three-dimensional
selectin models."
RL Biochem. Biophys. Res. Commun. 216:1018-1023(1995)
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- DATABASE: NAME-PROW; NOTE=CD guide CD62L entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62l.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25280; AAC63053.1; -
DR EMBL: X16150; CAB43475.1; -
DR EMBL: X17519; CAB43536.1; -
DR EMBL: X17519; CAB43537.1; ALT_SEQ.
DR EMBL: X16070; CAB34203.1; ALT_SEQ.
DR EMBL: M32414; AAB60700.1; -
DR EMBL: M32406; AAB60700.1; JOINED.
DR EMBL: M32407; AAB60700.1; JOINED.
DR EMBL: M32408; AAB60700.1; JOINED.
DR EMBL: M32409; AAB60700.1; JOINED.
DR EMBL: M32410; AAB60700.1; JOINED.
DR EMBL: M32411; AAB60700.1; JOINED.
DR EMBL: M32412; AAB60700.1; JOINED.
DR EMBL: M32413; AAB60700.1; JOINED.
DR PIR: A33912; A33912.
DR PIR: A34013; A34013.
DR PIR: S06798; S06798.
DR PIR: J10104; J10104.
DR PDB: 1KJB; 03-APR-96.
DR GLYCOSITEDB: P14151; -
DR MTM: 153240; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_C.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_C_1.
DR Pfam: PF00084; sushi_2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP_2.
DR SMART: SM00034; CLECT_1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00022; EGF_2_1.
DR PROSITE: PS00186; EGF_2_1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2_1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat; 3D-structure.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 33 355
FT DOMAIN 356 372
FT DOMAIN 355 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
L-SELECTIN.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
C-TYPE LECTIN (SHORT FORM).
EGF-LIKE.
SUSHI 1.
SUSHI 2.
BY SIMILARITY.

```

```

FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 37 37 D -> Y (IN REF. 4).
FT CONFLICT 178 178 Y -> H (IN REF. 4).
FT CONFLICT 193 193 F -> L (IN REF. 1 AND 4).
FT CONFLICT 213 213 P -> S (IN REF. 3).
FT CONFLICT 214 214 L -> F (IN REF. 4).
FT CONFLICT 218 220 SFS -> NFN (IN REF. 2).
FT CONFLICT 242 242 G -> E (IN REF. 2).
SQ SEQUENCE 372 AA; 42187 MW; 6EA9918ECA2D3643 CRC64;

```

```

Query Match 95.3%; Score 205; DB 1; Length 372;
Best Local Similarity 97.0%; Pred. No. 1, 8e-18;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CQPMSCGHECEVEIINNHTCNDVGYGPOCQ 33
Db 160 CQPMSCGHECEVEIINNHTCNDVGYGPOCQ 192

```

```

RESULT 2
LEML_MACMU STANDARD; PRT; 372 AA.
AC Q95198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (LECAM1) (CD62L).
GN SEIL.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsurushita N.;
RU Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U73730; AAB18246.1; -

```

```
DR HSP: P14151; IKJB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sushi_SCR_CCP.  
DR InterPro: IPR001304; Lectin_c.  
DR Pfam: PF00008; EGF; 1.  
DR Pfam: PF00059; lectin_c; 1.  
DR Pfam: PF00084; sushi; 2.  
DR PRINTS: PR00343; SELECTIN.  
DR SMART: SM00032; CCP; 2.  
DR SMART: SM00034; CLECT; 1.  
DR PROSITE: PS00022; EGF_1; 1.  
DR PROSITE: PS01186; EGF_2; 1.  
DR PROSITE: PS00615; C-TYPE-LECTIN_1; 1.  
DR PROSITE: PS00041; C-TYPE-LECTIN_2; 1.  
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 28 BY SIMILARITY.  
FT PROPEP 29 38 BY SIMILARITY.  
FT CHAIN 39 372 L-SELECTIN.  
FT DOMAIN 39 332 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 333 355 POTENTIAL.  
FT DOMAIN 356 372 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).  
FT DOMAIN 156 192 EGF-LIKE.  
FT DOMAIN 196 255 SUSHI 1.  
FT DOMAIN 258 317 SUSHI 2.  
FT DISULFID 57 155 BY SIMILARITY.  
FT DISULFID 128 147 BY SIMILARITY.  
FT DISULFID 160 171 BY SIMILARITY.  
FT DISULFID 165 180 BY SIMILARITY.  
FT DISULFID 182 191 BY SIMILARITY.  
FT DISULFID 197 241 BY SIMILARITY.  
FT DISULFID 227 254 BY SIMILARITY.  
FT DISULFID 259 303 BY SIMILARITY.  
FT DISULFID 289 316 BY SIMILARITY.  
FT CARBOHD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 372 AA; 42109 MW; 086D7DD5AC349D6D CRC64;  
  
Query Match 95.3%; Score 205; DB 1; Length 372;  
Best Local Similarity 97.0%; Pred. No. 1.8e-18;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 COPMSCSGHGECEIINNHTCNCMDVGYGPOCO 33  
DB 160 COPMSCSGHGECEIINNHTCNCMDVGYGPOCO 192  
  
RESULT 3  
LEMI_PANTR STANDARD; PRT; 372 AA.  
AC 095237;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion  
molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)  
DE (LECAM1) (CD62L).  
GN SEIL.  
OS Pan troglodytes (Chimpanzee).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
OX NCBI_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.
```

```
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,  
RA Tsunashita N.;  
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
CC -I- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE  
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL  
CC VENULES IN PERIPHERAL LYMPH NODES.  
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -I- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -I- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@ib-sib.ch)  
CC -----  
CC EMBL: U73728; ABL18248.1; -.  
DR HSP: P14151; IKJB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sushi_SCR_CCP.  
DR InterPro: IPR001304; Lectin_c.  
DR Pfam: PF00008; EGF; 1.  
DR Pfam: PF00059; lectin_c; 1.  
DR Pfam: PF00084; sushi; 2.  
DR PRINTS: PR00343; SELECTIN.  
DR SMART: SM00032; CCP; 2.  
DR SMART: SM00034; CLECT; 1.  
DR SMART: SM00181; EGF; 1.  
DR PROSITE: PS00022; EGF_1; 1.  
DR PROSITE: PS01186; EGF_2; 1.  
DR PROSITE: PS00615; C-TYPE-LECTIN_1; 1.  
DR PROSITE: PS00041; C-TYPE-LECTIN_2; 1.  
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 28 BY SIMILARITY.  
FT PROPEP 29 38 BY SIMILARITY.  
FT CHAIN 39 372 L-SELECTIN.  
FT DOMAIN 39 332 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 333 355 POTENTIAL.  
FT DOMAIN 356 372 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).  
FT DOMAIN 156 192 EGF-LIKE.  
FT DOMAIN 196 255 SUSHI 1.  
FT DOMAIN 258 317 SUSHI 2.  
FT DISULFID 57 155 BY SIMILARITY.  
FT DISULFID 128 147 BY SIMILARITY.  
FT DISULFID 160 171 BY SIMILARITY.  
FT DISULFID 165 180 BY SIMILARITY.  
FT DISULFID 182 191 BY SIMILARITY.  
FT DISULFID 197 241 BY SIMILARITY.  
FT DISULFID 227 254 BY SIMILARITY.  
FT DISULFID 259 303 BY SIMILARITY.  
FT DISULFID 289 316 BY SIMILARITY.  
FT CARBOHD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 372 AA; 42188 MW; 6EA991802A2D3643 CRC64;  
  
Query Match 95.3%; Score 205; DB 1; Length 372;  
Best Local Similarity 97.0%; Pred. No. 1.8e-18;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 COPMSCSGHGECEIINNHTCNCMDVGYGPOCO 33
```

```

Db      160 CQWCSGSGHCEVEIINNTCNCVDGYGPOCQ 192
|||||
RESULT  4
LEML_PAPHA STANDARD: PRT: 372 AA.
AC 028768;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
DE SEL.
GN Papio hamadryas (Hamadryas baboon).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97128794; PubMed=8973334;
RA Tsunushita N., Fu H., Berg E.L.;
RT PCR cloning of the cDNA encoding baboon L-selectin.*;
RL Gene 181:219-220(1996).
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U52074; AABA0903.1; -
DR HSSP: P14151; 1KJB.
DR InterPro: IPR000561; EGF-1like.
DR InterPro: IPR0002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sush1; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sush1; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 182
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
BY SIMILARITY.

```

```

FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 372 AA; 42091 MW; 64E7BD5AC349D69 CRC64;

Query Match 95.3%; Score 205; DB 1; Length 372;
Best Local Similarity 97.0%; Pred. No. 1,8e-18;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQWCSGSGHCEVEIINNTCNCVDGYGPOCQ 33
|||||
Db 160 CQWCSGSGHCEVEIINNTCNCVDGYGPOCQ 192

RESULT  5
LEML_PONPY STANDARD: PRT: 372 AA.
AC 095235;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
DE SEL.
GN Pongo pygmaeus (Orangutan).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RA Buchan J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsunushita N.;
RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RL
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U73729; AAB18247.1; -
DR HSSP: P14151; 1KJB.
DR InterPro: IPR000561; EGF-1like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sush1; 2.

```

```

DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP: 2.
DR SMART: SM00034; CLECT: 1.
DR SMART: SM00181; EGF: 1.
DR PROSITE: PS00022; EGF_1: 1.
DR PROSITE: PS01186; EGF_2: 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1: 1.
DR PROSITE: PS00615; C-TYPE LECTIN_2: 1.
DR PROSITE: PS00615; C-TYPE LECTIN_3: 1.
KM Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
  Signal; Sush1; Repeat.
FT SIGNAL 1 28 BY SIMILARITY.
FT PROPEP 29 38 BY SIMILARITY.
FT CHAIN 39 372 L-SELECTIN.
FT TRANSMEM 39 332 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 333 355 POTENTIAL.
FT DOMAIN 356 372 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 156 192 EGF-LIKE.
FT DOMAIN 196 255 SUSH1 1.
FT DOMAIN 258 317 SUSH1 2.
FT DISULFID 57 155 BY SIMILARITY.
FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 42118 MW; 6517DD2213FF15E CRC64;

```

```

Query Match 95.3%; Score 205; DB 1; Length 372;
Best Local Similarity 97.0%; Pred. No. 1.8e-18;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 COPMSCSGHCEVEIINHTCNCVDVGYGPQCQ 33
    |||||
Db 160 COPMSCSGHCEVEIINHTCNCVDVGYGPQCQ 192

```

```

RESULT 6
ID LEM1_BOVIN STANDARD: PRT: 370 AA.
AC P98131;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion
  molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (LECAM1) (CD62L).
GN SELL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCB1_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92164727; PubMed=1371468;
RA Walcheck B., White M., Kurk S., Kishimoto T.K., Juttla M.A.;
RT "Characterization of the bovine peripheral lymph node homing
  receptor: a lectin cell adhesion molecule (LECAM).";
RL Eur. J. Immunol. 22:469-476(1992).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=94055053; PubMed=7694420;
RA Bosworth B.T., Dowbenko D., Shuster D.E., Harp J.A.;
RT "Bovine L-selectin: a peripheral lymphocyte homing receptor.";
RL Vet. Immunol. Immunopathol. 37:201-215(1993).
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
  OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
  VENULES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSH1 (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X62882; CAA44676.1; -.
CC HSSP: P14151; IKTJ.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sush1; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP: 2.
DR SMART: SM00034; CLECT: 1.
DR SMART: SM00181; EGF: 1.
DR PROSITE: PS00022; EGF_1: 1.
DR PROSITE: PS01186; EGF_2: 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1: 1.
DR PROSITE: PS00615; C-TYPE LECTIN_2: 1.
DR PROSITE: PS00615; C-TYPE LECTIN_3: 1.
KM Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
  Signal; Sush1; Repeat.
FT SIGNAL 1 28 POTENTIAL.
FT PROPEP 29 38 POTENTIAL.
FT CHAIN 39 370 L-SELECTIN.
FT TRANSMEM 39 333 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 334 354 POTENTIAL.
FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 156 192 EGF-LIKE.
FT DOMAIN 196 255 SUSH1 1.
FT DOMAIN 258 317 SUSH1 2.
FT DISULFID 57 155 BY SIMILARITY.
FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 229 303 BY SIMILARITY.
FT DISULFID 259 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 370 AA; 41971 MW; 92168B8116A93228 CRC64;

```

```

Query Match 88.8%; Score 191; DB 1; Length 370;

```


DT 01-JUL-1993 (Rel. 26, last sequence update)
 DT 01-FEB-1996 (Rel. 33, last annotation update)
 DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (LY-22) (Lymphocyte surface MEL-14 antigen)
 DE (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
 GN SEL1 OR LNH1 OR LY-22.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92329548; PubMed=1378303;
 RA Matenabe T., Song Y., Hirayama Y., Tamatani T., Kuida K., Miyasaka M.;
 RT "Sequence and expression of a rat cDNA for LECAM-1.";
 RL Biochim. Biophys. Acta 1131:321-324(1992).
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D10831; BAA01613.1; -
 DR PIR: S23936; S23936.
 DR HSSP: P14151; 1KJB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sush1_2.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP_2.
 DR SMART: SM00034; CLECT_1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
 DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin; Selectin; Signal; Sush1; Repeat.
 FT SIGNAL 1 28
 FT PROPEP 29 38
 FT CHAIN 39 372
 FT DOMAIN 39 332
 FT TRANSMEM 333 335
 FT DOMAIN 356 372
 FT DOMAIN 155 155
 FT DOMAIN 156 192
 FT DOMAIN 196 255
 FT DOMAIN 258 317
 FT DISULFID 37 155
 FT DISULFID 128 147
 FT DISULFID 160 171
 FT DISULFID 165 180
 FT DISULFID 182 191
 FT DISULFID 197 241
 FT DISULFID 227 254
 FT DISULFID 259 303
 FT DISULFID 289 316
 FT CARBOHYD 60
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 AA; 42441 MW; 3888AE0F1EAD191A CRC64;
 SQ SEQUENCE
 Query Match 74.0%; Score 159; DB 1; Length 372;
 Best Local Similarity 78.8%; Pred. No. 9e-13;
 Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 OY 1 COPWCSGHCCEVEIINHHNCNDVGYGPOQC 33
 DB 160 COPESCRHCEVEITNNHTICDPGYGPOQC 192
 RESULT 9
 ID LEM2_CANFA STANDARD; PRT; 611 AA.
 AC P33730;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, last sequence update)
 DT 01-FEB-1996 (Rel. 33, last annotation update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1) (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2) (CD62E).
 GN SEL1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OK NCBI_TaxID=9615;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=jugular vein;
 RA Manning A.M., Lane C.L., Auchampach J.A., Kukielka G.L.,
 RA Rosenbloom C.L., Anderson D.C.;
 RL Submitted (XX-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY ELAM-1 IS SIALLYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF POLYACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF GLYCOPROTEINS).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L23087; AAA30843.1; -
 DR HSSP: P16581; IESL.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sush1_6.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP_6.
 DR SMART: SM00034; CLECT_1.
 DR SMART: SM00181; EGF_2.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.

DR PROSITE: PS50041; C-TYPE LECTIN 2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT CHAIN 1 22
FT SIGNAL 1 22
FT DOMAIN 23 611
FT TRANSMEM 558 579
FT DOMAIN 580 611
FT DOMAIN 39 139
FT DOMAIN 140 176
FT DOMAIN 180 239
FT DOMAIN 242 301
FT DOMAIN 304 364
FT DOMAIN 367 427
FT DOMAIN 430 490
FT DOMAIN 493 549
FT DISULFID 41 139
FT DISULFID 112 131
FT DISULFID 144 155
FT DISULFID 149 164
FT DISULFID 166 175
FT DISULFID 181 225
FT DISULFID 211 238
FT DISULFID 243 287
FT DISULFID 273 300
FT DISULFID 305 350
FT DISULFID 336 363
FT DISULFID 368 413
FT DISULFID 399 426
FT DISULFID 431 476
FT DISULFID 462 489
FT DISULFID 494 535
FT DISULFID 521 548
FT CARBOHYD 26 26
FT CARBOHYD 161 161
FT CARBOHYD 204 204
FT CARBOHYD 266 266
FT CARBOHYD 313 313
FT CARBOHYD 333 333
FT CARBOHYD 528 528
SQ SEQUENCE 611 AA; 66314 MW; 35DA9E3DF225E4F6 CRC64;

Query Match 60.9%; Score 131; DB 1; Length 611;
Best Local Similarity 60.6%; Pred. No. 4.2e-09;
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 1 CQWSCSGHCEVEIINNHCNCVGYGPOCQ 33
DB 144 CPTSCSGHCEVEIINNHCNCVGYGPOCQ 176

RESULT 10
LEM2_HUMAN STANDARD; PRT; 610 AA.
ID LEM2_HUMAN
AC P16581; P16111;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE OR ELAM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175359; PubMed=1689848;
RA Hession C., Osborn L., Goff D., Chi-Rosso G., Vassallo C.,
RA Pasek M., Pittack C., Tizard R., Goelz S., McCarthy K., Hopple S.,
RA Lobb R.;

RT RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning
RT and functional interactions.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).
RX [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=89162047; PubMed=2466335;
RX Bevilacqua M.P., Stengelin S., Gimbrone M.A. Jr., Seed B.;
RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for
RT neutrophils related to complement regulatory proteins and lectins.";
RT Science 243:1160-1165(1989).
RN [3]
RN SEQUENCE FROM N.A.
RN MEDLINE=91115870; PubMed=1703529;
RN Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T.,
RA Gimbrone M.A. Jr., Bevilacqua M.P.;
RT "Structure and chromosomal location of the gene for endothelial-
RT leukocyte adhesion molecule 1.";
RT J. Biol. Chem. 266:2466-2473(1991).
RN [4]
RN LIGAND.
RN MEDLINE=91068005; PubMed=1701274;
RN Phillips M.L., Nudelman E., Gaeta F.C., Perez M., Singhal A.K.,
RA Hakomori S., Paulson J.C.;
RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate
RT ligand, sialyl-Lex.";
RT Science 250:1130-1132(1990).
RN [5]
RN 3D-STRUCTURE MODELING OF LECTIN DOMAIN.
RN MEDLINE=93202275; PubMed=7681016;
RN Mills A.;
RT "Modelling the carbohydrate recognition domain of human E-selectin.";
RT FEBS Lett. 319:5-11(1993).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.
RN MEDLINE=94150646; PubMed=7509040;
RN Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S.,
RA Huang K.-S., Presky D.H., Familletti P.C., Wollitzky B.A., Burns D.K.;
RT "Insight into E-selectin/ligand interaction from the crystal
RT structure and mutagenesis of the lec/EGF domains.";
RT Nature 367:532-538(1994).
RN [7]
RN VARIANT ARG-149.
RN MEDLINE=95179107; PubMed=7533025;
RN Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S.,
RA Schulte K.L., Glaser C., Rohde K., Baumann G., Speer A.;
RT "E-selectin polymorphism and atherosclerosis: an association study.";
RT Hum. Mol. Genet. 3:1935-1937(1994).
RN [8]
RN VARIANT ARG-149.
RN MEDLINE=99134508; PubMed=9933738;
RN Ye S.-Q., Usher D., Virgil D., Zhang L.-Q., Yochim S.E., Gupta R.;
RT "A PstI polymorphism detects the mutation of serine-128 to arginine in
RT CD 62e gene - a risk factor for coronary artery disease.";
RT J. Biomed. Sci. 6:18-21(1999).
RN [9]
RN FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
RN MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
RN ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
RN POLYACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
RN GLYCOLIPIDS).
RN [10]
RN SUBCELLULAR LOCATION: Type I membrane protein.
RN [11]
RN POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A
RN HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY
RN HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH
RN ANGIOGRAPHICALLY PROVEN SEVERE ATHEROSCLEROSIS COMPARED WITH AN
RN UNSELECTED POPULATION (SER-149).
RN [12]
RN SIMILARITY: TO OTHER SELECTINS/LUCAS.
RN [13]
RN SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
RN [14]
RN SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
RN [15]
RN DATABASE: NAME=PROW; NOTE=CD guide CD62E entry;
RN WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62e.htm"
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M30640; AAA52377.1; -
 DR EMBL: M61893; AAA52375.1; -
 DR EMBL: M61895; AAA52375.1; JOINED.
 DR EMBL: M61887; AAA52375.1; JOINED.
 DR EMBL: M61888; AAA52375.1; JOINED.
 DR EMBL: M61890; AAA52375.1; JOINED.
 DR EMBL: M61891; AAA52375.1; JOINED.
 DR EMBL: M61892; AAA52375.1; JOINED.
 DR EMBL: M24736; AAA52376.1; -
 DR PIR: A32606; A32606.
 DR PIR: A35046; A35046.
 DR PIR: A38615; A38615.
 DR PDB: 1ESL; 31-AUG-94.
 DR PDB: 1KJA; 03-APR-96.
 DR MIM: 131210; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sushi; 6.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP; 6.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF_1; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C_Type-LECTIN_1; 1.
 DR PROSITE: PS00615; C_Type-LECTIN_2; 1.
 DR Cell adhesion: Transmembrane; Glycoprotein; EGF-like domain; lectin;
 KW Selectin; Signal; Sushi; Repeat; Polymorphism; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 610
 FT DOMAIN 22 556
 FT TRANSMEM 557 578
 FT DOMAIN 579 610
 FT DOMAIN 138 138
 FT DOMAIN 139 175
 FT DOMAIN 179 238
 FT DOMAIN 241 300
 FT DOMAIN 303 363
 FT DOMAIN 363 426
 FT DOMAIN 429 489
 FT DOMAIN 492 548
 FT DOMAIN 548 610
 FT DISULFID 40 138
 FT DISULFID 111 130
 FT DISULFID 143 154
 FT DISULFID 148 163
 FT DISULFID 165 174
 FT DISULFID 180 224
 FT DISULFID 210 237
 FT DISULFID 242 286
 FT DISULFID 272 299
 FT DISULFID 304 349
 FT DISULFID 335 362
 FT DISULFID 367 412
 FT DISULFID 398 425
 FT DISULFID 430 475
 FT DISULFID 461 488
 FT DISULFID 493 534
 FT DISULFID 520 547
 FT CARBOHYD 25 25
 FT CARBOHYD 145 145
 FT CARBOHYD 160 160
 FT CARBOHYD 179 179
 E-SELECTIN. (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 C-TYPE LECTIN (SHORT FORM).
 EGF-LIKE.
 SUSHI 1.
 SUSHI 2.
 SUSHI 3.
 SUSHI 4.
 SUSHI 5.
 SUSHI 6.

FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 149 149 S -> R (ASSOCIATED WITH A RISK FACTOR FOR CAD).
 FT VARIANT 295 295 /FTID-VAR_004191.
 FT VARIANT 421 421 E -> K (IN DBSNP:5364).
 FT VARIANT 468 468 /FTID-VAR_011791.
 FT VARIANT 468 468 E -> O (IN DBSNP:5366).
 FT VARIANT 575 575 /FTID-VAR_011792.
 FT VARIANT 575 575 H -> Y (IN DBSNP:5368).
 FT VARIANT 575 575 /FTID-VAR_011793.
 FT VARIANT 575 575 L -> F (IN DBSNP:5355).
 FT VARIANT 575 575 /FTID-VAR_011794.
 FT SEQUENCE 610 AA; 66655 MW; 7D43E3C0D1229229 CRC64;
 Query Match 60.5%; Score 130; DB 1; Length 610;
 Best local Similarity 63.6%; Pred. No. 5; 6e-09;
 Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 OY 1 COPWCSGHCCEVETINNTNCNDVGYGPOCQ 33
 DB 143 CFTNCSGHCCEVETINNTYCKDPPGSLKCE 175
 RESULT 11
 ID LEM2_RAT STANDARD; PRT; 549 AA.
 AC P98105;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 2) (LECAM2)
 DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
 DE (CD62E).
 GN SELE OR ELAM-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING ON LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC ELAM-1 IS SIAYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOPROTEINS).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L25527; AAA41113.1; -
 DR HSP; P16581; IKJA.

```

DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sushi_5.
DR SMART: SM00032; CCP; 5.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 1 21 E-SELECTIN.
FT DOMAIN 22 549 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 516 POTENTIAL.
FT DOMAIN 517 549 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 38 138 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 139 175 EGF-LIKE.
FT DOMAIN 179 239 SUSHI 1.
FT DOMAIN 242 301 SUSHI 2.
FT DOMAIN 304 364 SUSHI 3.
FT DOMAIN 367 427 SUSHI 4.
FT DOMAIN 430 486 SUSHI 5.
FT DISULFID 40 138 BY SIMILARITY.
FT DISULFID 111 130 BY SIMILARITY.
FT DISULFID 143 154 BY SIMILARITY.
FT DISULFID 148 163 BY SIMILARITY.
FT DISULFID 165 174 BY SIMILARITY.
FT DISULFID 180 225 BY SIMILARITY.
FT DISULFID 210 238 BY SIMILARITY.
FT DISULFID 243 287 BY SIMILARITY.
FT DISULFID 273 300 BY SIMILARITY.
FT DISULFID 305 350 BY SIMILARITY.
FT DISULFID 336 363 BY SIMILARITY.
FT DISULFID 368 413 BY SIMILARITY.
FT DISULFID 399 426 BY SIMILARITY.
FT DISULFID 431 472 BY SIMILARITY.
FT DISULFID 458 485 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SQUENCE 549 AA; 60079 MW; 85CEBCDB7B0144C8 CRC64;

```

```

Query Match 59.1%; Score 127; DB 1; Length 549;
Best Local Similarity 62.5%; Pred. No. 1, 2e-08;
Matches 20; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 1 CQPMSCGHEGCEVITNNHNCNDVGYGPOC 32
DB 143 CTTNCSGHEGCEVITNNHNCNDVGYGPOC 174

RESULT 12
LEM3_BOVIN STANDARD; PRT; 646 AA.
AC P42201;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)

```

```

DE (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN SEMP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Capillary endothelium;
RX MEDLINE-93249394; PubMed=7683458;
RA Strubel N.A., Nguyen M., Kansas G.S., Tedder T.F., Bischoff J.;
RT "Isolation and characterization of a bovine cDNA encoding a
functional homolog of human P-selectin."
RL Blochm. Biophys. Res. Commun. 192:338-344(1993)
CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS
CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC LEUCOCYTES. THE LIGAND RECOGNIZED IS SIATYL-LEWIS X.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
CC THE CELL SURFACE.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS; BOVINE P-LECTIN LACKS
CC THE HUMAN SUSHI-3, -4 AND -7 EQUIVALENTS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L12041; AAA30743.1; -.
DR HSSP: P16109; 1P5B.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sushi_5.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 6.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 41 BY SIMILARITY.
FT CHAIN 1 41 P-SELECTIN.
FT DOMAIN 42 646 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 588 611 POTENTIAL.
FT DOMAIN 612 646 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 58 158 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 159 195 EGF-LIKE.
FT DOMAIN 199 258 SUSHI 1.
FT DOMAIN 261 320 SUSHI 2.
FT DOMAIN 323 382 SUSHI 3.
FT DOMAIN 385 444 SUSHI 4.
FT DOMAIN 457 516 SUSHI 5.
FT DOMAIN 519 578 SUSHI 6.
FT DISULFID 60 158 BY SIMILARITY.
FT DISULFID 131 150 BY SIMILARITY.
FT DISULFID 163 174 BY SIMILARITY.

```

```
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 200 244 BY SIMILARITY.
FT DISULFID 230 257 BY SIMILARITY.
FT DISULFID 262 306 BY SIMILARITY.
FT DISULFID 292 319 BY SIMILARITY.
FT DISULFID 324 368 BY SIMILARITY.
FT DISULFID 354 381 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
FT DISULFID 416 443 BY SIMILARITY.
FT DISULFID 458 502 BY SIMILARITY.
FT DISULFID 488 515 BY SIMILARITY.
FT DISULFID 520 564 BY SIMILARITY.
FT DISULFID 550 577 BY SIMILARITY.
FT CARBOHYD 48 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 634 637 ENDOCYTOSIS SIGNAL (PROBABLE).
SO SEQUENCE 646 AA: 71229 MW: 573912A4627A6ACA CRC64;
```

```
Query Match 59.1%; Score 127; DB 1; Length 646;
Best Local Similarity 57.6%; Pred. No. 1.4e-08;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGEVEIINHTCNCNDVGYRPOCO 33
DB 163 CODMSCSGKQECIETIGNTCSCYPGFYRPECE 195
```

```
RESULT 13
LEM3_SHEEP STANDARD; PRT: 769 AA.
ID LEM3_SHEEP
AC P98109;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P-selectin precursor (Granule membrane protein 140) (CMP-140) (PADGEM)
DE (CD62p) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN SELP.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Burns S.A., Neufeld E.J., Donady J.J.;
RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L34270; AAB59261.1; -.
DR HSSP: P16109; IFSB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushi; 8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 32
FT CHAIN 33 769
FT DOMAIN 33 717
FT TRANSMEM 718 734
FT DOMAIN 735 759
FT DOMAIN 58 158
FT DOMAIN 159 195
FT DOMAIN 199 258
FT DOMAIN 261 320
FT DOMAIN 323 382
FT DOMAIN 385 444
FT DOMAIN 447 506
FT DOMAIN 509 568
FT DOMAIN 580 639
FT DOMAIN 642 701
FT DOMAIN 703 759
FT DISULFID 131 158
FT DISULFID 163 174
FT DISULFID 168 183
FT DISULFID 185 194
FT DISULFID 200 244
FT DISULFID 230 257
FT DISULFID 262 306
FT DISULFID 292 319
FT DISULFID 324 368
FT DISULFID 381 430
FT DISULFID 386 443
FT DISULFID 416 443
FT DISULFID 448 492
FT DISULFID 478 505
FT DISULFID 510 554
FT DISULFID 540 567
FT DISULFID 581 625
FT DISULFID 611 638
FT DISULFID 643 687
FT DISULFID 673 700
FT CARBOHYD 54 54
FT CARBOHYD 80 80
FT CARBOHYD 180 180
FT CARBOHYD 212 212
FT CARBOHYD 219 219
FT CARBOHYD 292 292
FT CARBOHYD 347 347
FT CARBOHYD 398 398
FT CARBOHYD 604 604
FT CARBOHYD 655 655
FT CARBOHYD 662 662
FT CARBOHYD 680 680
FT SITE 757 760
FT VARIANT 556 566
FT VARIANT 579 579
SO SEQUENCE 769 AA: 84317 MW: 23E42575D60FAB15 CRC64;
```


FT	DISULFID	144	155	BY SIMILARITY.
FT <th>DISULFID</th> <th>149</th> <th>164</th> <th>BY SIMILARITY.</th>	DISULFID	149	164	BY SIMILARITY.
FT <th>DISULFID</th> <th>166</th> <th>175</th> <th>BY SIMILARITY.</th>	DISULFID	166	175	BY SIMILARITY.
FT <th>DISULFID</th> <th>181</th> <th>224</th> <th>BY SIMILARITY.</th>	DISULFID	181	224	BY SIMILARITY.
FT <th>DISULFID</th> <th>210</th> <th>237</th> <th>BY SIMILARITY.</th>	DISULFID	210	237	BY SIMILARITY.
FT <th>DISULFID</th> <th>242</th> <th>286</th> <th>BY SIMILARITY.</th>	DISULFID	242	286	BY SIMILARITY.
FT <th>DISULFID</th> <th>272</th> <th>299</th> <th>BY SIMILARITY.</th>	DISULFID	272	299	BY SIMILARITY.
FT <th>DISULFID</th> <th>304</th> <th>349</th> <th>BY SIMILARITY.</th>	DISULFID	304	349	BY SIMILARITY.
FT <th>DISULFID</th> <th>335</th> <th>362</th> <th>BY SIMILARITY.</th>	DISULFID	335	362	BY SIMILARITY.
FT <th>DISULFID</th> <th>367</th> <th>408</th> <th>BY SIMILARITY.</th>	DISULFID	367	408	BY SIMILARITY.
FT <th>DISULFID</th> <th>394</th> <th>421</th> <th>BY SIMILARITY.</th>	DISULFID	394	421	BY SIMILARITY.
FT <th>CARBOHYD</th> <th>61</th> <th>61</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	61	61	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>79</th> <th>79</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	79	79	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>88</th> <th>88</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	88	88	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>163</th> <th>161</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	163	161	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>203</th> <th>203</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	203	203	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>265</th> <th>265</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	265	265	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>312</th> <th>312</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	312	312	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>316</th> <th>316</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	316	316	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>379</th> <th>379</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	379	379	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <th>CARBOHYD</th> <th>401</th> <th>401</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL).</th>	CARBOHYD	401	401	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ <th>SEQUENCE</th> <th>485 AA;</th> <th>53200 MW;</th> <th>AE93IC9B521E3904 CRC64;</th>	SEQUENCE	485 AA;	53200 MW;	AE93IC9B521E3904 CRC64;
Query Match				
Best Local Similarity		54.9%;	Score 118;	DB 1;
Matches 18;		Conservative	Pred. No. 1.4e-07;	Length 485;
			4;	Mismatches 11;
				Indels 0;
				Gaps 0;
OY	1	COPWSCGSHGECVEIINHTCNDVCYVPCQC	33	
DB	144	CNPFCGSHGECVEITNNYTCQHFEGFKCE	176	
RESULT 19				
LEM2_RABIT	ID	LEM2_RABIT	STANDARD;	PRT; 551 AA.
AC	P27113;			
DT	01-AUG-1992	(Rel. 23, Created)		
DT	01-AUG-1992	(Rel. 23, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	E-selectin precursor (Endothelial leukocyte adhesion molecule 1)			
DE	(ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)			
DE	(CP62E).			
GN	SELE.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Heart;			
RX	MEDLINE=92189729; PubMed=1372169;			
RA	Larigan J.D., Tsang T.C., Rumberger J.M., Burns D.K.;			
RT	"Characterization of cDNA and genomic sequences encoding rabbit			
RT	ELAM-1: conservation of structure and functional interactions with			
RT	leukocytes.";			
RL	DNA Cell Biol. 11:149-162(1992).			
CC	-1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND			
CC	MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY			
CC	ELAM-1 IS STALY-LWIS X (ALPHA(1-3)FUCOSYLATED DERIVATIVES OF			
CC	POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF			
CC	GLYCOLIPIDS).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- INDUCTION: BY CYTOKINES.			
CC	-1- SIMILARITY: TO OTHER SELECTINS/LECAMS.			
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
FT DOMAIN 159 195 EGF-LIKE.
FT DOMAIN 199 258 SUSH1 1.
FT DOMAIN 261 320 SUSH1 2.
FT DOMAIN 323 382 SUSH1 3.
FT DOMAIN 385 444 SUSH1 4.
FT DOMAIN 447 506 SUSH1 5.
FT DOMAIN 509 568 SUSH1 6.
FT DOMAIN 579 638 SUSH1 7.
FT DOMAIN 641 700 SUSH1 8.
FT DISULFID 60 158 BY SIMILARITY.
FT DISULFID 131 150 BY SIMILARITY.
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 200 244 BY SIMILARITY.
FT DISULFID 230 257 BY SIMILARITY.
FT DISULFID 262 306 BY SIMILARITY.
FT DISULFID 292 319 BY SIMILARITY.
FT DISULFID 324 368 BY SIMILARITY.
FT DISULFID 354 381 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
FT DISULFID 416 443 BY SIMILARITY.
FT DISULFID 448 492 BY SIMILARITY.
FT DISULFID 478 505 BY SIMILARITY.
FT DISULFID 510 554 BY SIMILARITY.
FT DISULFID 540 567 BY SIMILARITY.
FT DISULFID 580 624 BY SIMILARITY.
FT DISULFID 610 637 BY SIMILARITY.
FT DISULFID 642 686 BY SIMILARITY.
FT DISULFID 672 699 BY SIMILARITY.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 654 654 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 661 661 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 745 745 PALMITATE (BY SIMILARITY).
FT SITE 756 759 ENDOCYTOSIS SIGNAL (PROBABLE).
SQ SEQUENCE 768 AA: 83517 MW: 26FD7E8A5F3F1316 CRC64;
```

Query Match 50.7%; Score 109; DB 1; Length 768;
Best Local Similarity 48.5%; Pred. No. 2.8e-06;
Matches 16; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINHTCNCVGYGPOCQ 33
DB 163 CQMSGSGERIEFTIGTSCYCPGYEPECE 195

RESULT 21
ID FA12_CAVPO STANDARD; PRT: 603 AA.
AC 004962;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC TISSUE=Liver;
RA MEDLINE=93003367; PubMed=1390917;
RA Sema U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
Kambata T., Okabe H.;

```
RT "Primary structure of guinea-pig Hageman factor: sequence around the  
RT cleavage site differs from the human molecule.";  
RL Biochim. Biophys. Acta 1159:113-121(1992).  
CC -I- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN  
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE  
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.  
CC -I- CATALYTIC ACTIVITY: Cleaves selectively Arg-1-ile bonds in factor  
CC VII to form factor VIIa and factor XI to form factor XIa.  
CC -I- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HKK KININOGEN FORM A  
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY  
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST  
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR  
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIa.  
CC -I- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.  
CC -I- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.  
CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: X68615; CAA48600.1; -.  
DR HSSP: P00750; IRTF.  
DR MEROPS: S01.211; -.  
DR InterPro: IPR000561; EGF-1-like.  
DR InterPro: IPR000742; EGF 2.  
DR InterPro: IPR001881; EGF Ca.  
DR InterPro: IPR000562; FN_Type_II.  
DR InterPro: IPR000001; Kringie.  
DR InterPro: IPR001254; Trypsin.  
DR InterPro: IPR000083; fibronectin_type_1.  
DR Pfam: PF00008; EGF_2.  
DR Pfam: PF00039; fn1_1.  
DR Pfam: PF00040; fn2_1.  
DR Pfam: PF00051; kringie_1.  
DR Pfam: PF00089; trypsin_1.  
DR ProDom: PD000995; FN_Type_II; 1.  
DR SMART: SM00181; EGF_2.  
DR SMART: SM00058; FN1; 1.  
DR SMART: SM00059; FN2; 1.  
DR SMART: SM00130; KR; 1.  
DR SMART: SM00020; tryp_Spec; 1.  
DR PROSITE: PS00022; EGF_1; 2.  
DR PROSITE: PS01186; EGF_2; 1.  
DR PROSITE: PS01253; FIBRONECTIN_1; 1.  
DR PROSITE: PS00023; FIBRONECTIN_2; 1.  
DR PROSITE: PS00021; KRINGIE_1; 1.  
DR PROSITE: PS00070; KRINGIE_2; 1.  
DR PROSITE: PS50240; TRYPSIN_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN_SER; 1.  
KW Glycoprotein; Blood coagulation; Plasma; Kringie; Serine protease;  
KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
FT NON_TER 1 1  
FT SIGNAL <1 18  
FT CHAIN 19 358 ALPHA-FACTOR X11A HEAVY CHAIN.  
FT CHAIN 359 603 ALPHA-FACTOR X11A LIGHT CHAIN.  
FT DOMAIN 46 87 FIBRONECTIN TYPE-II.  
FT DOMAIN 93 130 EGF-LIKE 1.  
FT DOMAIN 132 172 FIBRONECTIN TYPE-I.  
FT DOMAIN 173 209 EGF-LIKE 2.  
FT DOMAIN 216 294 KRINGIE.  
FT DOMAIN 312 342 PRO-RICH.  
FT DOMAIN 359 603 SERINE PROTEASE.  
FT ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT_SITE 447 447 CHANGE RELAY SYSTEM (BY SIMILARITY).
```

```
FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 118 BY SIMILARITY.
FT DISULFID 120 129 BY SIMILARITY.
FT DISULFID 134 162 BY SIMILARITY.
FT DISULFID 160 169 BY SIMILARITY.
FT DISULFID 177 188 BY SIMILARITY.
FT DISULFID 182 197 BY SIMILARITY.
FT DISULFID 199 208 BY SIMILARITY.
FT DISULFID 216 294 BY SIMILARITY.
FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9BD59 CRC64;

Query Match 44.2%; Score 95; DB 1; Length 603;
Best Local Similarity 44.1%; Pred. No. 0.00012;
Matches 15; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCGHGECVEITNNHTCNCNDVGYGPOCOL 34
DB 177 CQTNPLNGRCLEVEGHHLCDCPKMYTPFCDL 210

RESULT 22
DILI_HUMAN STANDARD; PRT; 723 AA.
AC 000548;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Delta-like protein 1 precursor (Drosophila Delta homolog 1) (Delta1)
DE (H-Delta-1).
GN DILI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBITaxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.L.,
RA Banks A., Lehman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
RT "Human ligands of the Notch receptor.";
RL Am. J. Pathol. 154:785-794(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Oda T., Chandrasekharappa S.C.;
RT "Human Delta 1 gene sequence.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
CC MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
CC UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF003522; AAB61286.1; -.
DR EMBL; AF222310; AAG09716.1; -.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001774; DSL.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 8.
DR PRINTS; PRO0010; EGFBLD.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 1.
DR Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
KW Differentiation; Glycoprotein.
FT SIGNAL 1 17
FT CHAIN 18 723
FT DOMAIN 18 545 DELTA-LIKE PROTEIN 1.
FT TRANSMEM 546 568 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 569 723 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 226 254 EGF-LIKE 1.
FT DOMAIN 257 285 EGF-LIKE 2.
FT DOMAIN 292 325 EGF-LIKE 3.
FT DOMAIN 332 363 EGF-LIKE 4.
FT DOMAIN 370 402 EGF-LIKE 5.
FT DOMAIN 409 440 EGF-LIKE 6.
FT DOMAIN 447 478 EGF-LIKE 7.
FT DOMAIN 485 516 EGF-LIKE 8.
FT DOMAIN 517 548 EGF-LIKE 9.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 230 243 BY SIMILARITY.
FT DISULFID 245 254 BY SIMILARITY.
FT DISULFID 257 268 BY SIMILARITY.
FT DISULFID 263 274 BY SIMILARITY.
FT DISULFID 276 285 BY SIMILARITY.
FT DISULFID 292 304 BY SIMILARITY.
FT DISULFID 298 314 BY SIMILARITY.
FT DISULFID 316 325 BY SIMILARITY.
FT DISULFID 332 343 BY SIMILARITY.
FT DISULFID 337 352 BY SIMILARITY.
FT DISULFID 354 363 BY SIMILARITY.
FT DISULFID 370 381 BY SIMILARITY.
FT DISULFID 375 391 BY SIMILARITY.
FT DISULFID 393 402 BY SIMILARITY.
FT DISULFID 409 420 BY SIMILARITY.
FT DISULFID 414 429 BY SIMILARITY.
FT DISULFID 431 440 BY SIMILARITY.
FT DISULFID 447 467 BY SIMILARITY.
FT DISULFID 469 478 BY SIMILARITY.
FT DISULFID 485 496 BY SIMILARITY.
FT DISULFID 490 505 BY SIMILARITY.
FT DISULFID 507 516 BY SIMILARITY.
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 723 AA; 77956 MW; B4EC455FEFA32A12B CRC64;

Query Match 43.3%; Score 93; DB 1; Length 723;
Best Local Similarity 38.2%; Pred. No. 0.00026;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCGHGECVEITNNHTCNCNDVGYGPOCOL 34
DB 332 CDPSPCKNGSGCDLNSYSCTCPDPFGYKICEL 365
```

RESULT 23
NOTC_XENLA STANDARD; PRT; 2524 AA.
ID NOTC_XENLA
AC P21783;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurogenic locus notch protein homologue precursor (XOTCH protein).
GN XOTCH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_Taxid=8335;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=90385285; PubMed=2402639;
RA Coffman C., Harris W., Kintner C.;
RT "Xotch, the Xenopus homolog of Drosophila notch.";
RL Science, 249:1438-1441(1990).
RM [2]
RP REVISIONS TO 1759-1782.
RA Kintner C.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1 SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1 SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1 SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1 SIMILARITY: CONTAINS 6 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; M33874; AAB02039.1; -.
DR PIR; A35844; A35844.
DR HSSP; P00740; 1EDM.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR000800; Notch.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 36.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR01452; NOTCH.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF_CA; 23.
DR SMART; SM00001; EGF-like; 11.
DR SMART; SM00004; NL; 2.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 23.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 29.
DR PROSITE; PS01187; EGF_CA; 21.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2524 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG.
FT DOMAIN 20 1728 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1729 1730 POTENTIAL.
FT DOMAIN 1751 2524 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 57 EGF-LIKE 1.
FT DOMAIN 58 99 EGF-LIKE 2.

FT DOMAIN 102 140 EGF-LIKE 3.
FT DOMAIN 141 177 EGF-LIKE 4.
FT DOMAIN 179 215 EGF-LIKE 5.
FT DOMAIN 217 254 EGF-LIKE 6.
FT DOMAIN 256 292 EGF-LIKE 7.
FT DOMAIN 294 332 EGF-LIKE 8.
FT DOMAIN 334 370 EGF-LIKE 9.
FT DOMAIN 371 409 EGF-LIKE 10.
FT DOMAIN 411 449 EGF-LIKE 11.
FT DOMAIN 451 487 EGF-LIKE 12.
FT DOMAIN 489 525 EGF-LIKE 13.
FT DOMAIN 527 563 EGF-LIKE 14.
FT DOMAIN 565 600 EGF-LIKE 15.
FT DOMAIN 602 638 EGF-LIKE 16.
FT DOMAIN 640 675 EGF-LIKE 17.
FT DOMAIN 677 713 EGF-LIKE 18.
FT DOMAIN 715 750 EGF-LIKE 19.
FT DOMAIN 752 788 EGF-LIKE 20.
FT DOMAIN 790 826 EGF-LIKE 21.
FT DOMAIN 828 866 EGF-LIKE 22.
FT DOMAIN 868 904 EGF-LIKE 23.
FT DOMAIN 906 942 EGF-LIKE 24.
FT DOMAIN 944 980 EGF-LIKE 25.
FT DOMAIN 982 1018 EGF-LIKE 26.
FT DOMAIN 1020 1056 EGF-LIKE 27.
FT DOMAIN 1058 1094 EGF-LIKE 28.
FT DOMAIN 1096 1142 EGF-LIKE 29.
FT DOMAIN 1144 1180 EGF-LIKE 30.
FT DOMAIN 1182 1218 EGF-LIKE 31.
FT DOMAIN 1220 1264 EGF-LIKE 32.
FT DOMAIN 1266 1304 EGF-LIKE 33.
FT DOMAIN 1306 1346 EGF-LIKE 34.
FT DOMAIN 1347 1383 EGF-LIKE 35.
FT DOMAIN 1386 1424 EGF-LIKE 36.
FT REPEAT 1441 1478 LIN/NOTCH 1.
FT REPEAT 1479 1520 LIN/NOTCH 2.
FT REPEAT 1521 1560 LIN/NOTCH 3.
FT REPEAT 1561 1599 ANK 1.
FT REPEAT 1599 1624 ANK 2.
FT REPEAT 1624 1649 ANK 3.
FT REPEAT 1649 1674 ANK 4.
FT REPEAT 1674 1699 ANK 5.
FT REPEAT 1699 1724 ANK 6.
FT REPEAT 1724 1750 BY SIMILARITY.
FT DISULFID 22 35 BY SIMILARITY.
FT DISULFID 29 45 BY SIMILARITY.
FT DISULFID 47 56 BY SIMILARITY.
FT DISULFID 62 74 BY SIMILARITY.
FT DISULFID 68 87 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 111 128 BY SIMILARITY.
FT DISULFID 130 139 BY SIMILARITY.
FT DISULFID 145 156 BY SIMILARITY.
FT DISULFID 150 165 BY SIMILARITY.
FT DISULFID 167 176 BY SIMILARITY.
FT DISULFID 183 194 BY SIMILARITY.
FT DISULFID 188 203 BY SIMILARITY.
FT DISULFID 205 214 BY SIMILARITY.
FT DISULFID 221 232 BY SIMILARITY.
FT DISULFID 226 242 BY SIMILARITY.
FT DISULFID 244 253 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY.
FT DISULFID 265 280 BY SIMILARITY.
FT DISULFID 282 291 BY SIMILARITY.
FT DISULFID 298 311 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY.
FT DISULFID 322 331 BY SIMILARITY.
FT DISULFID 338 349 BY SIMILARITY.
FT DISULFID 343 358 BY SIMILARITY.
FT DISULFID 360 369 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY.
FT DISULFID 399 408 BY SIMILARITY.

```

FT DISULFID 415 428 BY SIMILARITY.
FT DISULFID 422 437 BY SIMILARITY.
FT DISULFID 439 448 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 460 475 BY SIMILARITY.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 504 BY SIMILARITY.
FT DISULFID 498 513 BY SIMILARITY.
FT DISULFID 515 524 BY SIMILARITY.
FT DISULFID 531 542 BY SIMILARITY.
FT DISULFID 536 551 BY SIMILARITY.
FT DISULFID 553 562 BY SIMILARITY.
FT DISULFID 569 579 BY SIMILARITY.
FT DISULFID 574 588 BY SIMILARITY.
FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 617 BY SIMILARITY.
FT DISULFID 611 626 BY SIMILARITY.
FT DISULFID 628 637 BY SIMILARITY.
FT DISULFID 644 654 BY SIMILARITY.
FT DISULFID 649 663 BY SIMILARITY.
FT DISULFID 665 674 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
FT DISULFID 686 701 BY SIMILARITY.
FT DISULFID 703 712 BY SIMILARITY.
FT DISULFID 719 729 BY SIMILARITY.
FT DISULFID 724 738 BY SIMILARITY.
FT DISULFID 740 749 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 761 776 BY SIMILARITY.
FT DISULFID 778 787 BY SIMILARITY.
FT DISULFID 794 805 BY SIMILARITY.
FT DISULFID 799 814 BY SIMILARITY.
FT DISULFID 816 825 BY SIMILARITY.
FT DISULFID 832 843 BY SIMILARITY.
FT DISULFID 837 854 BY SIMILARITY.
FT DISULFID 856 865 BY SIMILARITY.
FT DISULFID 872 883 BY SIMILARITY.
FT DISULFID 877 892 BY SIMILARITY.
FT DISULFID 894 903 BY SIMILARITY.
FT DISULFID 910 921 BY SIMILARITY.
FT DISULFID 915 930 BY SIMILARITY.
FT DISULFID 932 941 BY SIMILARITY.
FT DISULFID 986 997 BY SIMILARITY.
FT DISULFID 991 1006 BY SIMILARITY.
FT DISULFID 1008 1017 BY SIMILARITY.
FT DISULFID 1024 1035 BY SIMILARITY.
FT DISULFID 1029 1044 BY SIMILARITY.
FT DISULFID 1046 1055 BY SIMILARITY.
FT DISULFID 1062 1073 BY SIMILARITY.
FT DISULFID 1067 1082 BY SIMILARITY.
FT DISULFID 1084 1093 BY SIMILARITY.
FT DISULFID 1100 1121 BY SIMILARITY.
FT DISULFID 1115 1130 BY SIMILARITY.
FT DISULFID 1132 1141 BY SIMILARITY.
FT DISULFID 1148 1159 BY SIMILARITY.
FT DISULFID 1153 1168 BY SIMILARITY.

```

Query Match Best Local Similarity 43.3%; Score 93; DB 1; Length 2524;
 Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSHGCEVEITNNHTCNDGYVGPCCO 33
 DB 910 COPNPGHNGSCSDGIMFPCNCPAGFRRKCE 942

RESULT 24
 YNXX_CABEL STANDARD; PRT: 3051 AA.
 AC P34576;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

DE Hypothetical protein T20G5.3 in chromosome III (Fragment).
GN T20G5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=BRISTOL N2.
RA Berns M., Smith A.;
RL Submitted (Mar-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 45 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 SEA DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 VMFA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z30423; CAA83007.1; -.
CC PIR: S42373; S42373.
CC HSSP: P02468; 1TLE.
DR WormPep: T20G5.3; CE00478.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000082; SEA.
DR InterPro: IPR002035; VMFA.
DR Pfam: PF00008; EGF: 30.
DR Pfam: PF01390; SEA: 2.
DR Pfam: PF00092; vwa: 1.
DR SMART: SM00179; EGF_Ca: 1.
DR SMART: SM00001; EGF_Like: 38.
DR SMART: SM00200; SEA: 2.
DR SMART: SM00327; vwa: 1.
DR PROSITE: PS00010; ASX_HYDROXYL: 28.
DR PROSITE: PS00022; EGF_1: 1.
DR PROSITE: PS01186; EGF_2: 5.
DR PROSITE: PS01187; EGF_Ca: 1.
DR PROSITE: PS50024; SEA: 2.
DR PROSITE: PS50234; VMFA: 1.
KW Hypothetical protein; EGF-like domain; Repeat; Transmembrane.
FT NON_TER 1
FT DOMAIN <1 2701
FT TRANSMEM 2702 2722
FT DOMAIN 2723 3051
FT DOMAIN <1 44
FT DOMAIN 46 94
FT DOMAIN 100 144
FT DOMAIN 145 192
FT DOMAIN 194 245
FT DOMAIN 247 296
FT DOMAIN 313 354
FT DOMAIN 355 402
FT DOMAIN 404 452
FT DOMAIN 454 503
FT DOMAIN 514 690
FT DOMAIN 705 750
FT DOMAIN 750 794
FT DOMAIN 805 846
FT DOMAIN 847 892
FT DOMAIN 892 940
FT DOMAIN 942 990
FT DOMAIN 992 1039
FT DOMAIN 1043 1091
FT DOMAIN 1093 1144
FT DOMAIN 1146 1195
FT DOMAIN 1197 1245
FT DOMAIN 1247 1295

```

FT DOMAIN 1298 1346 EGF-LIKE 23.
FT DOMAIN 1352 1396 EGF-LIKE 24.
FT DOMAIN 1397 1444 EGF-LIKE 25.
FT DOMAIN 1446 1492 EGF-LIKE 26.
FT DOMAIN 1494 1542 EGF-LIKE 27.
FT DOMAIN 1544 1592 EGF-LIKE 28.
FT DOMAIN 1594 1642 EGF-LIKE 29.
FT DOMAIN 1644 1692 EGF-LIKE 30.
FT DOMAIN 1693 1739 EGF-LIKE 31.
FT DOMAIN 1740 1788 EGF-LIKE 32.
FT DOMAIN 1797 1847 EGF-LIKE 33.
FT DOMAIN 1849 1900 EGF-LIKE 34.
FT DOMAIN 1902 1950 EGF-LIKE 35.
FT DOMAIN 1952 1998 EGF-LIKE 36.
FT DOMAIN 2000 2047 EGF-LIKE 37.
FT DOMAIN 2047 2095 EGF-LIKE 38.
FT DOMAIN 2117 2156 EGF-LIKE 39.
FT DOMAIN 2157 2281 SEA 1.
FT DOMAIN 2283 2332 EGF-LIKE 40.
FT DOMAIN 2333 2456 SEA 2.
FT DOMAIN 2460 2504 EGF-LIKE 41.
FT DOMAIN 2508 2556 EGF-LIKE 42.
FT DOMAIN 2556 2608 EGF-LIKE 43.
FT DOMAIN 2612 2657 EGF-LIKE 44.
FT DOMAIN 2657 2693 EGF-LIKE 45.
FT DISULFID 2 13 BY SIMILARITY.
FT DISULFID 7 22 BY SIMILARITY.
FT DISULFID 24 43 BY SIMILARITY.
FT DISULFID 50 63 BY SIMILARITY.
FT DISULFID 57 72 BY SIMILARITY.
FT DISULFID 74 93 BY SIMILARITY.
FT DISULFID 104 120 BY SIMILARITY.
FT DISULFID 112 129 BY SIMILARITY.
FT DISULFID 131 143 BY SIMILARITY.
FT DISULFID 149 163 BY SIMILARITY.
FT DISULFID 157 172 BY SIMILARITY.
FT DISULFID 174 191 BY SIMILARITY.
FT DISULFID 198 214 BY SIMILARITY.
FT DISULFID 208 223 BY SIMILARITY.
FT DISULFID 225 244 BY SIMILARITY.
FT DISULFID 251 265 BY SIMILARITY.
FT DISULFID 259 274 BY SIMILARITY.
FT DISULFID 276 295 BY SIMILARITY.
FT DISULFID 317 330 BY SIMILARITY.
FT DISULFID 324 339 BY SIMILARITY.
FT DISULFID 341 353 BY SIMILARITY.
FT DISULFID 359 371 BY SIMILARITY.
FT DISULFID 365 380 BY SIMILARITY.
FT DISULFID 382 401 BY SIMILARITY.
FT DISULFID 408 421 BY SIMILARITY.
FT DISULFID 415 430 BY SIMILARITY.
FT DISULFID 432 451 BY SIMILARITY.
FT DISULFID 458 472 BY SIMILARITY.
FT DISULFID 466 481 BY SIMILARITY.
FT DISULFID 483 502 BY SIMILARITY.
FT DISULFID 709 725 BY SIMILARITY.
FT DISULFID 717 734 BY SIMILARITY.
FT DISULFID 736 749 BY SIMILARITY.
FT DISULFID 754 768 BY SIMILARITY.
FT DISULFID 762 778 BY SIMILARITY.
FT DISULFID 780 793 BY SIMILARITY.
FT DISULFID 809 822 BY SIMILARITY.
FT DISULFID 816 831 BY SIMILARITY.
FT DISULFID 833 845 BY SIMILARITY.
FT DISULFID 851 867 BY SIMILARITY.
FT DISULFID 876 891 BY SIMILARITY.
FT DISULFID 878 891 BY SIMILARITY.
FT DISULFID 896 909 BY SIMILARITY.
FT DISULFID 903 918 BY SIMILARITY.
FT DISULFID 920 939 BY SIMILARITY.
FT DISULFID 946 959 BY SIMILARITY.
FT DISULFID 953 968 BY SIMILARITY.
FT DISULFID 970 989 BY SIMILARITY.

CALCIUM-BINDING (POTENTIAL).

FT DISULFID 996 1010 BY SIMILARITY.
FT DISULFID 1004 1019 BY SIMILARITY.
FT DISULFID 1021 1038 BY SIMILARITY.
FT DISULFID 1047 1060 BY SIMILARITY.
FT DISULFID 1054 1070 BY SIMILARITY.
FT DISULFID 1072 1090 BY SIMILARITY.
FT DISULFID 1097 1113 BY SIMILARITY.
FT DISULFID 1105 1122 BY SIMILARITY.
FT DISULFID 1124 1143 BY SIMILARITY.
FT DISULFID 1150 1164 BY SIMILARITY.
FT DISULFID 1157 1173 BY SIMILARITY.
FT DISULFID 1175 1194 BY SIMILARITY.
FT DISULFID 1201 1214 BY SIMILARITY.
FT DISULFID 1208 1223 BY SIMILARITY.
FT DISULFID 1225 1244 BY SIMILARITY.
FT DISULFID 1251 1264 BY SIMILARITY.
FT DISULFID 1258 1273 BY SIMILARITY.
FT DISULFID 1275 1294 BY SIMILARITY.
FT DISULFID 1302 1315 BY SIMILARITY.
FT DISULFID 1309 1324 BY SIMILARITY.
FT DISULFID 1326 1345 BY SIMILARITY.
FT DISULFID 1356 1372 BY SIMILARITY.
FT DISULFID 1364 1381 BY SIMILARITY.
FT DISULFID 1383 1395 BY SIMILARITY.
FT DISULFID 1401 1415 BY SIMILARITY.
FT DISULFID 1409 1424 BY SIMILARITY.
FT DISULFID 1426 1443 BY SIMILARITY.
FT DISULFID 1450 1464 BY SIMILARITY.
FT DISULFID 1458 1473 BY SIMILARITY.
FT DISULFID 1475 1491 BY SIMILARITY.
FT DISULFID 1498 1512 BY SIMILARITY.
FT DISULFID 1506 1521 BY SIMILARITY.
FT DISULFID 1523 1541 BY SIMILARITY.
FT DISULFID 1548 1562 BY SIMILARITY.
FT DISULFID 1556 1571 BY SIMILARITY.
FT DISULFID 1573 1591 BY SIMILARITY.
FT DISULFID 1598 1611 BY SIMILARITY.
FT DISULFID 1605 1620 BY SIMILARITY.
FT DISULFID 1622 1641 BY SIMILARITY.
FT DISULFID 1648 1661 BY SIMILARITY.
FT DISULFID 1655 1670 BY SIMILARITY.
FT DISULFID 1672 1691 BY SIMILARITY.
FT DISULFID 1697 1709 BY SIMILARITY.
FT DISULFID 1703 1719 BY SIMILARITY.
FT DISULFID 1721 1738 BY SIMILARITY.
FT DISULFID 1744 1758 BY SIMILARITY.
FT DISULFID 1752 1767 BY SIMILARITY.
FT DISULFID 1769 1787 BY SIMILARITY.
FT DISULFID 1801 1815 BY SIMILARITY.

Query Match 41.9%; Score 90; DB 1; Length 3051;

Best local Similarity 51.6%; Pred. No. 0.0023;

Matches 16; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 6 CSGHGCEV--IINHTCNDVGYGPOCOL 34

Db 2664 CNAHGDCVHTARNNTTCVCTDGMWGPOCOV 2694

RESULT 25
ID PRTC_MOUSE STANDARD; PRT; 461 AA.
AC P33587; O35498;
DT 01-FEB-1994 (Rel. 28. Created)
DT 01-FEB-1994 (Rel. 28. Last sequence update)
DT 16-OCT-2001 (Rel. 40. Last annotation update)
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE (Autoproteolysin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV).
DE PROC.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=9231697; PubMed=1618739;
RA Tada N., Sato M., Tsujimura A., Iwase R., Hashimoto-Gotoh T.;
RT "Isolation and characterization of a mouse protein C cDNA.";
RL J. Biochem. 111:491-495(1992).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RA MEDLINE=98152576; PubMed=9493582;
RA Jalbert L.R., Rosen E.D., Lissens A., Carmeliet P., Collen D.,
RA Castellino F.J.;
RT "Nucleotide structure and characterization of the murine gene encoding
RT anticoagulant protein C.";
RL Thromb. Haemost. 79:310-316(1998).
RN [3]
RN SEQUENCE OF 274-434 FROM N.A.
RC STRAIN=BALB/C;
RA MEDLINE=94318474; PubMed=8043441;
RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT "A comparative study of partial primary structures of the catalytic
RT region of mammalian protein C.";
RL Br. J. Haematol. 86:590-600(1994).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and VIIIA.
CC -1- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC STRONGLY PROMOTED BY THROMBOMODULIN.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- PTM: THE VITAMIN K-DEPENDENT ENZYMIC CARBOXYLATION OF SOME
CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
CC SITE IS NECESSARY FOR THE RECOGNITION OF THE
CC THROMBIN-THROMBOMODULIN COMPLEX.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRIPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10445; BAA01235.1; -;
DR EMBL: AF034569; AAC33795.1; -;
DR EMBL: D43755; BAA07812.1; -;
DR PIR: JX0210; JX0210.
DR HSSP: P04070; 1PCU.
DR MEROPS: S01.218; -;
DR MGD: MGI:97771; PROC.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR002383; GLA_blood.
DR InterPro: IPR001254; Trypsin.
DR InterPro: IPR000294; Vitk_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PRO0722; CHYMOTRYPSIN.

DR PRINTS: PRO0001; GLABLOOD.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM00001; EGF_Like; 1.
DR SMART: SM00069; GLA; 1.
DR SMART: SM00020; TRYP_Spc; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Blood coagulation: Glycoprotein; Serine protease;
KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT SIGNAL 1 33
FT PROPEP 34 41
FT CHAIN 42 196
FT CHAIN 199 461
FT PEPTIDE 199 212
FT SITE 212 213
FT DOMAIN 96 131
FT DOMAIN 135 175
FT DOMAIN 213 461
FT MOD_RES 47 47
FT MOD_RES 48 48
FT MOD_RES 48 48
FT MOD_RES 55 55
FT MOD_RES 57 57
FT MOD_RES 60 60
FT MOD_RES 61 61
FT MOD_RES 66 66
FT MOD_RES 67 67
FT MOD_RES 70 70
FT MOD_RES 70 70
FT MOD_RES 112 112
FT ACT_SITE 253 253
FT ACT_SITE 299 299
FT ACT_SITE 402 402
FT DISULFID 58 63
FT DISULFID 91 110
FT DISULFID 100 105
FT DISULFID 104 119
FT DISULFID 121 130
FT DISULFID 139 150
FT DISULFID 146 159
FT DISULFID 161 174
FT DISULFID 162 319
FT DISULFID 238 254
FT DISULFID 373 387
FT DISULFID 398 426
FT CARBOHYD 214 290
FT CARBOHYD 290 355
FT CARBOHYD 355 355
FT CONFLICT 328 328
FT CONFLICT 393 393
SQ SEQUENCE 461 AA; 51945 MW; 53FAAD085194D6E CRC64;

Query Match Score 87; DB 1; Length 461;
Best Local Similarity 46.48; Pred. No. 0.00094;
Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
OY 6 CSGHGCEVELINNHGTCNCDVGYGQCQ 33
| | | | | : : : | : | | | : | |

DB 104 CCGHCTCIDIGSCSCDKMEKFCQ 131

RESULT 26

DLL1_RAT

ID DLL1_RAT STANDARD: PRT: 714 AA.

AC P97677;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Delta-like protein 1 precursor (Drosophila Delta homolog 1) (Delta).

GN DLL1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Disibio G., Hebshi L., Boulter J., Weimaster G.;

RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U78889; AAB37343.1; -

DR HSSP: P08709; I8P9; -

DR InterPro: IPR000152; Asx_hydroxyl.

DR InterPro: IPR001774; DSL.

DR InterPro: IPR000561; EGF-Like.

DR InterPro: IPR000742; EGF_2.

DR InterPro: IPR001881; EGF_Ca.

DR InterPro: IPR001438; EGF_IT.

DR Pfam: PF01414; DSL; 1.

DR Pfam: PF00008; EGF; 7.

DR PRINTS: PR00010; EGFBL00D.

DR SMART: SM00051; DSL; 1.

DR SMART: SM00179; EGF_CA; 4.

DR SMART: SM00001; EGF_Like; 4.

DR PROSITE: PS00010; ASX_HYDROXYL; 3.

DR PROSITE: PS00022; EGF_1; 8.

DR PROSITE: PS01186; EGF_2; 8.

DR PROSITE: PS01187; EGF_CA; 2.

KW Signal: EGF-like domain; Repeat: Transmembrane; Developmental protein; Differentiation: Glycoprotein.

KW STGNAL 1 17

FT CHAIN 1 714 DELTA-LIKE PROTEIN 1.

FT DOMAIN 18 537 EXTRACELLULAR (POTENTIAL).

FT TRANSLEM 538 560 POTENTIAL.

FT DOMAIN 561 714 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 225 253 EGF-LIKE 1.

FT DOMAIN 256 284 EGF-LIKE 2.

FT DOMAIN 291 324 EGF-LIKE 3.

FT DOMAIN 331 362 EGF-LIKE 4.

FT DOMAIN 369 401 EGF-LIKE 5.

FT DOMAIN 408 439 EGF-LIKE 6.

FT DOMAIN 446 477 EGF-LIKE 7.

FT DOMAIN 484 515 EGF-LIKE 8.

FT DISULFID 225 236 BY SIMILARITY.

FT DISULFID 229 242 BY SIMILARITY.

FT DISULFID 244 253 BY SIMILARITY.

FT DISULFID 256 267 BY SIMILARITY.

FT DISULFID 262 273 BY SIMILARITY.

FT DISULFID 275 284 BY SIMILARITY.

FT DISULFID 291 303 BY SIMILARITY.

FT DISULFID 297 313 BY SIMILARITY.

FT DISULFID 315 324 BY SIMILARITY.

FT DISULFID 331 342 BY SIMILARITY.

FT DISULFID 336 351 BY SIMILARITY.

FT DISULFID 353 362 BY SIMILARITY.

FT DISULFID 369 380 BY SIMILARITY.

FT DISULFID 374 390 BY SIMILARITY.

FT DISULFID 392 401 BY SIMILARITY.

FT DISULFID 408 419 BY SIMILARITY.

FT DISULFID 413 428 BY SIMILARITY.

FT DISULFID 430 439 BY SIMILARITY.

FT DISULFID 446 466 BY SIMILARITY.

FT DISULFID 468 477 BY SIMILARITY.

FT DISULFID 484 495 BY SIMILARITY.

FT DISULFID 489 504 BY SIMILARITY.

FT DISULFID 506 515 BY SIMILARITY.

FT CARBOHYD 476 476 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 714 AA; 77378 MW; 4B8E2272BAEA27E CRC64;

Query Match 40.5%; Score 87; DB 1; Length 714;

Best Local Similarity 35.3%; Pred. No. 0.0014;

Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

OY 1 CQPSGSGHGCVCIIINHTNCNDVGYGPQCL 34

DB 331 CAPSPCRNGSCYDLEDSYSTCPCPGYGRVCEL 364

RESULT 27

FBP1_STRPU

ID FBP1_STRPU STANDARD: PRT: 1064 AA.

AC P10079;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fibronectin I precursor (Epidermal growth factor-related protein 1) (UEGF-1).

GN EGF1.

OS Strongylocentrotus purpuratus (Purple sea urchin).

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoda; Strongylocentrotidae; CC Strongylocentrotus.

OX NCBI_TaxID=7668;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90112459; PubMed=2514273;

RA Delgado-Illio-Reynoso M.G., Rollo D.R., Hursh D.A., Raff R.A.;

RT "Structural analysis of the uegf gene in the sea urchin Strongylocentrotus purpuratus reveals more similarity to vertebrate than to invertebrate genes with EGF-like repeats."

RL J. Mol. Evol. 29:314-327(1989).

RN [2]

RP SEQUENCE OF 279-476 AND 781-1064 FROM N.A.

RX MEDLINE=87319677; PubMed=3498216;

RA Hursh D.A., Andrews M.E., Raff R.A.;

RT "A sea urchin gene encodes a polypeptide homologous to epidermal growth factor."

RL Science 237:1487-1490(1987).

RN [3]

RP AVIDIN-LIKE DOMAIN.

RX MEDLINE=89196806; PubMed=2784773;

RA Hunt L.T., Barker W.C.;

RT "Avidin-like domain in an epidermal growth factor homolog from a sea urchin."

RL FASEB J. 3:1760-1764(1989).

RN [4]

RP CHARACTERIZATION.

RX MEDLINE=91285254; PubMed=2060714;

RA Bisgrove B.W., Andrews M.E., Raff R.A.;

FT DISULFID 869 884 BY SIMILARITY.
FT DISULFID 886 895 BY SIMILARITY.
FT DISULFID 902 913 BY SIMILARITY.
FT DISULFID 907 922 BY SIMILARITY.
FT DISULFID 924 933 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 851 851 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 477 780 MISSING (IN ISOFORM IB).
FT CONFLICT 279 279 L -> S (IN REF. 2).
SQ SEQUENCE 1064 AA; 112072 MW; 2E569CA012ED6D09 CRC64;

Query Match 40.5%; Score 87; DB 1; Length 1064;
Best Local Similarity 39.4%; Pred. No. 0.002;
Matches 13; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

OY 1 CQPMSCSGHCEVEIINHTCCNDGYGPGQC 33
DB 750 CASMPCNLNGACIEVMNGYTCOCVAGYGVICE 782

RESULT 28
CRB_DROME
ID CRB_DROME STANDARD: PRT; 2139 AA.
AC P10040.
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Crumbs protein precursor (95F).
GN CRB.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R; TISSUE=Embryo;
RX MEDLINE=90263104; PubMed=2344615;
RA Tepass U., Theres C., Knust E.;
RT "Crumbs encodes an EGF-like protein expressed on apical membranes of
RT Drosophila epithelial cells and required for organization of
RT epithelia.";
RT Cell 61:787-799(1990).
RN [2]
RP SEQUENCE OF 1663-1955 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=87218537; PubMed=3107986;
RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D.,
RA Vaassin H., Campos-Ortega J.A.;
RT "Egf homologous sequences encoded in the genome of Drosophila
RT melanogaster, and their relation to neurogenic genes.";
RT EMBO J. 6:761-766(1987).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE DEVELOPMENT OF EPITHELIA,
CC POSSIBLY FOR THE ESTABLISHMENT AND/OR MAINTENANCE OF CELL
CC POLARITY. IT MAY ACT AS A SIGNAL.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- PTM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
CC
CC "THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M33753; AAA28428.1; ALT_SEQ.
CC EMBL: X05144; CAA28793.1; -
CC PIR: B26637; B26637.

DR PIR: A35672; A35672.
DR HSSP: P00740; 11XA.
DR FlyBase: FBgn0000368; crb.
DR InterPro: IPR000152; Asx_Hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF 2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_T1.
DR InterPro: IPR001791; Laminin_G.
DR Pfam: PF00008; EGF; 26.
DR Pfam: PF00054; Laminin_G; 3.
DR PRINTS: PR00010; EGFBL00D.
DR SMART: SM00179; EGF_CA; 11.
DR SMART: SM00001; EGF-like; 16.
DR SMART: SM00282; LamG; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 15.
DR PROSITE: PS00022; EGF_1; 26.
DR PROSITE: PS01186; EGF_2; 17.
DR PROSITE: PS01187; EGF_CA; 15.
DR PROSITE: PS50025; LAM_G_DOMAIN; 3.
KW Differentiation; Repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Signal; Phosphorylation.
FT SIGNAL 1 90
FT CHAIN 91 2139 CRUMBS PROTEIN.
FT DOMAIN 91 2084 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2085 2111 POTENTIAL.
FT DOMAIN 2112 2139 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 267 303 EGF-LIKE 1.
FT DOMAIN 306 343 EGF-LIKE 2.
FT DOMAIN 348 386 EGF-LIKE 3.
FT DOMAIN 388 425 EGF-LIKE 4.
FT DOMAIN 427 463 EGF-LIKE 5.
FT DOMAIN 464 500 EGF-LIKE 6.
FT DOMAIN 501 532 EGF-LIKE 7.
FT DOMAIN 545 581 EGF-LIKE 8.
FT DOMAIN 582 611 EGF-LIKE 9.
FT DOMAIN 609 646 EGF-LIKE 10.
FT DOMAIN 648 685 EGF-LIKE 11.
FT DOMAIN 687 723 EGF-LIKE 12.
FT DOMAIN 725 761 EGF-LIKE 13.
FT DOMAIN 763 800 EGF-LIKE 14.
FT DOMAIN 802 838 EGF-LIKE 15.
FT DOMAIN 840 902 EGF-LIKE 16.
FT DOMAIN 904 940 EGF-LIKE 17.
FT DOMAIN 942 978 EGF-LIKE 18.
FT DOMAIN 980 1021 EGF-LIKE 19.
FT DOMAIN 1023 1205 LAMININ G-LIKE 1.
FT DOMAIN 1207 1243 EGF-LIKE 20.
FT DOMAIN 1250 1480 LAMININ G-LIKE 2.
FT DOMAIN 1481 1517 EGF-LIKE 21.
FT DOMAIN 1558 1758 LAMININ G-LIKE 3.
FT DOMAIN 1759 1795 EGF-LIKE 22.
FT DOMAIN 1797 1833 EGF-LIKE 23.
FT DOMAIN 1835 1871 EGF-LIKE 24.
FT DOMAIN 1874 1915 EGF-LIKE 25.
FT DOMAIN 1915 1951 EGF-LIKE 26.
FT DOMAIN 1953 1989 EGF-LIKE 27.
FT DOMAIN 1991 2029 EGF-LIKE 28.
FT DOMAIN 2030 2070 EGF-LIKE 29.
FT DISULFID 276 282 BY SIMILARITY.
FT DISULFID 276 291 BY SIMILARITY.
FT DISULFID 293 302 BY SIMILARITY.
FT DISULFID 310 321 BY SIMILARITY.
FT DISULFID 315 331 BY SIMILARITY.
FT DISULFID 333 342 BY SIMILARITY.
FT DISULFID 352 363 BY SIMILARITY.
FT DISULFID 357 374 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 392 403 BY SIMILARITY.
FT DISULFID 397 412 BY SIMILARITY.
FT DISULFID 414 424 BY SIMILARITY.
FT DISULFID 431 442 BY SIMILARITY.
FT DISULFID 436 451 BY SIMILARITY.

```

FT DISULFID 453 462 BY SIMILARITY.
FT DISULFID 468 479 BY SIMILARITY.
FT DISULFID 473 488 BY SIMILARITY.
FT DISULFID 480 499 BY SIMILARITY.
FT DISULFID 505 515 BY SIMILARITY.
FT DISULFID 509 520 BY SIMILARITY.
FT DISULFID 522 531 BY SIMILARITY.
FT DISULFID 549 562 BY SIMILARITY.
FT DISULFID 556 569 BY SIMILARITY.
FT DISULFID 571 580 BY SIMILARITY.
FT DISULFID 586 597 BY SIMILARITY.
FT DISULFID 591 602 BY SIMILARITY.
FT DISULFID 604 610 BY SIMILARITY.
FT DISULFID 613 624 BY SIMILARITY.
FT DISULFID 618 634 BY SIMILARITY.
FT DISULFID 636 645 BY SIMILARITY.
FT DISULFID 652 664 BY SIMILARITY.
FT DISULFID 659 673 BY SIMILARITY.
FT DISULFID 675 684 BY SIMILARITY.
FT DISULFID 691 702 BY SIMILARITY.
FT DISULFID 696 711 BY SIMILARITY.
FT DISULFID 713 722 BY SIMILARITY.
FT DISULFID 729 740 BY SIMILARITY.
FT DISULFID 734 749 BY SIMILARITY.
FT DISULFID 751 760 BY SIMILARITY.
FT DISULFID 767 778 BY SIMILARITY.
FT DISULFID 772 787 BY SIMILARITY.
FT DISULFID 789 799 BY SIMILARITY.
FT DISULFID 806 817 BY SIMILARITY.
FT DISULFID 811 826 BY SIMILARITY.
FT DISULFID 828 837 BY SIMILARITY.
FT DISULFID 844 855 BY SIMILARITY.
FT DISULFID 849 880 BY SIMILARITY.
FT DISULFID 892 901 BY SIMILARITY.
FT DISULFID 908 919 BY SIMILARITY.
FT DISULFID 913 928 BY SIMILARITY.
FT DISULFID 930 939 BY SIMILARITY.
FT DISULFID 946 957 BY SIMILARITY.
FT DISULFID 952 966 BY SIMILARITY.
FT DISULFID 968 977 BY SIMILARITY.
FT DISULFID 984 995 BY SIMILARITY.
FT DISULFID 1009 1020 BY SIMILARITY.
FT DISULFID 1011 1020 BY SIMILARITY.
FT DISULFID 1211 1222 BY SIMILARITY.
FT DISULFID 1216 1231 BY SIMILARITY.
FT DISULFID 1233 1242 BY SIMILARITY.
FT DISULFID 1485 1496 BY SIMILARITY.
FT DISULFID 1490 1505 BY SIMILARITY.
FT DISULFID 1507 1516 BY SIMILARITY.
FT DISULFID 1563 1574 BY SIMILARITY.
FT DISULFID 1763 1774 BY SIMILARITY.
FT DISULFID 1768 1783 BY SIMILARITY.
FT DISULFID 1785 1794 BY SIMILARITY.
FT DISULFID 1801 1812 BY SIMILARITY.
FT DISULFID 1806 1821 BY SIMILARITY.
FT DISULFID 1833 1850 BY SIMILARITY.
FT DISULFID 1839 1852 BY SIMILARITY.
FT DISULFID 1844 1859 BY SIMILARITY.
FT DISULFID 1861 1870 BY SIMILARITY.
FT DISULFID 1878 1889 BY SIMILARITY.
FT DISULFID 1883 1903 BY SIMILARITY.
FT DISULFID 1905 1914 BY SIMILARITY.
FT DISULFID 1919 1930 BY SIMILARITY.
FT DISULFID 1924 1939 BY SIMILARITY.
FT DISULFID 1941 1950 BY SIMILARITY.
FT DISULFID 1957 1968 BY SIMILARITY.
FT DISULFID 1962 1977 BY SIMILARITY.
FT DISULFID 1979 1988 BY SIMILARITY.
FT DISULFID 1995 2008 BY SIMILARITY.
FT DISULFID 2002 2017 BY SIMILARITY.
FT DISULFID 2028 2037 BY SIMILARITY.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 40.0%; Score 86; DB 1; Length 2139;
Best Local Similarity 38.2%; Pred. No. 0.0053;
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Qy 1 COPWSCSGHCEVEIINNHTCNDVGYGPOCOL 34
Db 908 CDSNPSCKHGNCNDIGTGYTCECEPGEETHCEI 941

RESULT 29
NOTC_DROME STANDARD; PRT; 2703 AA.
AC P07207; P04154; 097458; 09W478;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neurogenic locus Notch protein precursor.
GN N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R; TISSUE=Embryo;
RA MEDLINE=86079359; PubMed=3935325;
RT Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
RT "Nucleotide sequence from the neurogenic locus notch implies a gene
RT product that shares homology with proteins containing EGF-like
RT repeats."
RN Cell 43:567-581(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Canton-S; and Oregon-R; TISSUE=Embryo;
RA MEDLINE=87064624; PubMed=3097517;
RT Kidd S., Kelley M.R., Young M.W.;
RT "Sequence of the notch locus of Drosophila melanogaster: relationship
RT of the encoded protein to mammalian clotting and growth factors."
RN Mol. Cell. Biol. 6:3094-3108(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Burks K.C., Busman D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```


Best Local Similarity 42.4%; Pred. No. 0.0064;
Matches 14; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

OY 1 CQPMSCSGHGEVETIINNHTCNCVDVGYGPOCQ 33

Db 1026 CRESSCNGSCIDINGNCISGLAGSGANCO 1058

RESULT 30

FP2_MYTGA STANDARD: PRT: 473 AA.
ID FP2_MYTGA Q25464;
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adhesive plaque matrix protein 2 precursor (Foot protein 2) (MGFP2)
DE (MGFP-2).
GN FP2.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Foot;
RX MEDLINE=95204464; PubMed=7896812;
RA Inoue K., Takeuchi Y., Miki D., Odo S.;
RT "Mussel adhesive plaque protein gene is a novel member of epidermal growth factor-like gene family."
RL J. Biol. Chem. 270:6698-6701(1995).
CC -1- FUNCTION: PROVIDES ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSELS PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A FIBROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION STARTS AT THE PEDIVELIGER, FOOT FORMATION, STAGE.
CC -1- PTM: SOME TYROSINES ARE HYDROXYLATED (THUS PRODUCING DOPA - 3,4-DIHYDROXYPHENYLALANINE).
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL: DA3794; BAA07852.1; -.
DR HSSP; P00740; IEDM.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF_11.
DR SMART; SM00181; EGF_10.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; 11.
DR PROSITE; PS01186; EGF_2; 10.
KW Glycoprotein; EGF-like domain; Repeat; Signal; Hydroxylation.
FT SIGNAL 1 17
FT CHAIN 18 473
FT DOMAIN 45 81
FT DOMAIN 82 117
FT DOMAIN 118 154
FT DOMAIN 155 191
FT DOMAIN 192 228
FT DOMAIN 229 265
FT DOMAIN 266 301
FT DOMAIN 302 340
FT DOMAIN 342 378
FT DOMAIN ADHESIVE PLAUQUE MATRIX PROTEIN 2.
FT EGF-LIKE 1.
FT EGF-LIKE 2.
FT EGF-LIKE 3.
FT EGF-LIKE 4.
FT EGF-LIKE 5.
FT EGF-LIKE 6.
FT EGF-LIKE 7.
FT EGF-LIKE 8.
FT EGF-LIKE 9.

FT DOMAIN 383 420
FT DOMAIN 425 461
FT MOD_RES 23
FT MOD_RES 31 31
FT MOD_RES 36 36
FT MOD_RES 43 43
FT CARBOHYD 93 93
FT CARBOHYD 49 60
FT DISULFID 54 69
FT DISULFID 71 80
FT DISULFID 86 97
FT DISULFID 91 106
FT DISULFID 108 117
FT DISULFID 122 133
FT DISULFID 127 143
FT DISULFID 145 154
FT DISULFID 159 170
FT DISULFID 164 180
FT DISULFID 182 191
FT DISULFID 196 207
FT DISULFID 201 217
FT DISULFID 219 228
FT DISULFID 233 244
FT DISULFID 238 254
FT DISULFID 256 265
FT DISULFID 270 281
FT DISULFID 275 290
FT DISULFID 292 301
FT DISULFID 306 317
FT DISULFID 311 328
FT DISULFID 330 339
FT DISULFID 346 357
FT DISULFID 351 366
FT DISULFID 368 377
FT DISULFID 387 399
FT DISULFID 393 408
FT DISULFID 410 419
FT DISULFID 429 440
FT DISULFID 434 449
FT DISULFID 451 460
SQ SEQUENCE 473 AA; 51772 MW; BA76BA8C3BA9A0F CRC64;

Query Match 39.8%; Score 85.5; DB 1; Length 473;
Best Local Similarity 38.2%; Pred. No. 0.0015;
Matches 13; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

OY 1 CQPMSCSGHGEVETIINNHTCNCVDVGYGPOCQ 33

Db 196 CKPNCNGSGTCSADKFGDYSCCRPGFPGPCE 229

RESULT 31
TENA_HUMAN STANDARD: PRT: 2201 AA.
ID TENA_HUMAN P24821; Q15567; Q14583;
AC 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neurexectin) (GWM) (J1) (Mitochondrius antigen) (Glioma-associated-extracellular matrix antigen) (GP 150-225) (Tenascin-C) (TN-C).
GN HXB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-32.
RC TISSUE=Fetal brain, and Melanoma;
RX MEDLINE=91187670; PubMed=1707164;
RA Sirtl A., Carmemolla B., Saginatti M., Leprini A., Casari G., Baralle F., Zardi L.,


```
FT DISULFID 456 465 BY SIMILARITY.
FT DISULFID 470 480 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 487 496 BY SIMILARITY.
FT DISULFID 501 511 BY SIMILARITY.
FT DISULFID 505 516 BY SIMILARITY.
FT DISULFID 518 527 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 536 547 BY SIMILARITY.
FT DISULFID 549 558 BY SIMILARITY.
FT DISULFID 563 573 BY SIMILARITY.
FT DISULFID 567 578 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 594 604 BY SIMILARITY.
FT DISULFID 598 609 BY SIMILARITY.
FT DISULFID 611 620 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1034 1034 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1079 1079 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1093 1093 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1119 1119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1184 1184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1210 1210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1301 1301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1366 1366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1392 1392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1455 1455 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
Query Match 39.1% Score 84; DB 1; Length 2201;
Best Local Similarity 53.3% Pred. No. 0.0094;
Matches 16; Conservative 0; Mismatches 10; Indels 4; Gaps 1;
```

```
Qy 3 PWSGSGHGCEVETINHTGNCVGYGPOC 32
Db 409 PNGSGHGRCV---NGQCVDEGYTGEDC 434
```

```
RESULT 32
NTC3_MOUSE STANDARD: PRT: 2318 AA.
ID 061982;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurogenic locus notch 3 protein.
GN NOTCH3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X SWISS WEBSTER;
RX MEDLINE=95001556; PubMed=7918097;
RA Lardelli M., Dalstrand J., Lendahl U.;
RT "The novel Notch homologue mouse Notch 3 lacks specific epidermal
RT growth factor-repeats and is expressed in proliferating
RT neuroepithelium.";
RL Mech. Dev. 46:123-136(1994).
CC -1- FUNCTION: NOTCH 1, 2 AND 3 PLAY A COMBINATIONAL ROLE DURING
CC VARIOUS CELL FATE DECISIONS AND MORPHOLOGICAL MOVEMENTS IN THE
CC DEVELOPING CNS AND PROBABLY OTHER REGIONS OF THE EMBRYO.
CC -1- TISSUE SPECIFICITY: PROLIFERATING NEUROEPITHELIUM.
CC -1- DEVELOPMENTAL STAGE: CNS DEVELOPMENT.
CC -1- SIMILARITY: CONTAINS 34 EGF-LIKE DOMAINS.
```

```
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X74760; CAA52776.1; -.
CC HSSP: P00740; IIXA.
CC MGD: MGI:99460; Notch3.
CC InterPro: IPR002110; ANK.
CC InterPro: IPR000152; Asx_hydroxyl.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR000742; EGF-2.
CC InterPro: IPR001881; EGF_Ca.
CC InterPro: IPR001438; EGF-II.
CC InterPro: IPR000800; Notch.
CC Pfam: PF00023; ank; 6.
CC Pfam: PF00008; EGF; 34.
CC Pfam: PF00066; notch; 3.
CC PRINTS: PR00010; EGFLOOD.
CC PRINTS: PR01452; NOTCH.
CC SMART: SM00248; ANK; 5.
CC SMART: SM00179; EGF_CA; 19.
CC SMART: SM00001; EGF_Like; 13.
CC SMART: SM00004; NL; 3.
CC PROSITE: PS50088; ANK_REPEAT; 4.
CC PROSITE: PS50097; ANK_REPEAT_REGION; 1.
CC PROSITE: PS00022; EGF_1; 33.
CC PROSITE: PS01186; EGF_2; 27.
CC PROSITE: PS01187; EGF_Ca; 17.
CC DR Differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane;
CC KW ANK repeat; Glycoprotein.
CC KM
CC FT DOMAIN 1 1643
CC FT TRANSMEM 1644 1664
CC FT DOMAIN 1665 2318
CC FT DOMAIN 2242 2261
CC FT DOMAIN 39 78
CC FT DOMAIN 79 119
CC FT DOMAIN 120 157
CC FT DOMAIN 159 196
CC FT DOMAIN 198 235
CC FT DOMAIN 237 273
CC FT DOMAIN 275 313
CC FT DOMAIN 315 351
CC FT DOMAIN 352 380
CC FT DOMAIN 392 430
CC FT DOMAIN 432 468
CC FT DOMAIN 470 506
CC FT DOMAIN 508 544
CC FT DOMAIN 546 581
CC FT DOMAIN 583 619
CC FT DOMAIN 621 656
CC FT DOMAIN 658 694
CC FT DOMAIN 696 731
CC FT DOMAIN 735 771
CC FT DOMAIN 772 809
CC FT DOMAIN 811 848
CC FT DOMAIN 850 886
CC FT DOMAIN 888 923
CC FT DOMAIN 925 961
CC FT DOMAIN 963 999
CC FT DOMAIN 1001 1035
CC FT DOMAIN 1037 1083
CC FT DOMAIN 1085 1121
CC FT DOMAIN 1123 1159
CC FT DOMAIN 1161 1204
CC FT DOMAIN 1206 1245
CC -----
CC EXTRACELLULAR.
CC POTENTIAL.
CC CYTOPLASMIC.
CC PEST.
CC EGF-LIKE 1.
CC EGF-LIKE 2.
CC EGF-LIKE 3.
CC EGF-LIKE 4.
CC EGF-LIKE 5.
CC EGF-LIKE 6.
CC EGF-LIKE 7.
CC EGF-LIKE 8.
CC EGF-LIKE 9.
CC EGF-LIKE 10.
CC EGF-LIKE 11.
CC EGF-LIKE 12.
CC EGF-LIKE 13.
CC EGF-LIKE 14.
CC EGF-LIKE 15.
CC EGF-LIKE 16.
CC EGF-LIKE 17.
CC EGF-LIKE 18.
CC EGF-LIKE 19.
CC EGF-LIKE 20.
CC EGF-LIKE 21.
CC EGF-LIKE 22.
CC EGF-LIKE 23.
CC EGF-LIKE 24.
CC EGF-LIKE 25.
CC EGF-LIKE 26.
CC EGF-LIKE 27.
CC EGF-LIKE 28.
CC EGF-LIKE 29.
CC EGF-LIKE 30.
CC EGF-LIKE 31.
CC EGF-LIKE 32.
```

FT DOMAIN 1247 1288 EGF-LIKE 32.
FT DOMAIN 1290 1326 EGF-LIKE 33.
FT DOMAIN 1336 1374 EGF-LIKE 34.
FT REPEAT 1388 1428 LIN/NOTCH 1.
FT REPEAT 1429 1467 LIN/NOTCH 2.
FT REPEAT 1468 1503 LIN/NOTCH 3.
FT REPEAT 1839 1868 ANK 1.
FT REPEAT 1872 1902 ANK 2.
FT REPEAT 1906 1935 ANK 3.
FT REPEAT 1939 1968 ANK 4.
FT REPEAT 1972 2001 ANK 5.
FT DISULFID 43 55 BY SIMILARITY.
FT DISULFID 49 66 BY SIMILARITY.
FT DISULFID 63 77 BY SIMILARITY.
FT DISULFID 88 94 BY SIMILARITY.
FT DISULFID 109 118 BY SIMILARITY.
FT DISULFID 124 135 BY SIMILARITY.
FT DISULFID 129 145 BY SIMILARITY.
FT DISULFID 147 156 BY SIMILARITY.
FT DISULFID 163 175 BY SIMILARITY.
FT DISULFID 169 184 BY SIMILARITY.
FT DISULFID 186 195 BY SIMILARITY.
FT DISULFID 202 213 BY SIMILARITY.
FT DISULFID 207 223 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 241 252 BY SIMILARITY.
FT DISULFID 246 261 BY SIMILARITY.
FT DISULFID 263 272 BY SIMILARITY.
FT DISULFID 279 292 BY SIMILARITY.
FT DISULFID 286 301 BY SIMILARITY.
FT DISULFID 303 312 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 324 339 BY SIMILARITY.
FT DISULFID 341 350 BY SIMILARITY.
FT DISULFID 356 367 BY SIMILARITY.
FT DISULFID 361 378 BY SIMILARITY.
FT DISULFID 380 389 BY SIMILARITY.
FT DISULFID 396 409 BY SIMILARITY.
FT DISULFID 403 418 BY SIMILARITY.
FT DISULFID 420 429 BY SIMILARITY.
FT DISULFID 436 447 BY SIMILARITY.
FT DISULFID 441 456 BY SIMILARITY.
FT DISULFID 458 467 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 479 494 BY SIMILARITY.
FT DISULFID 496 505 BY SIMILARITY.
FT DISULFID 512 523 BY SIMILARITY.
FT DISULFID 517 532 BY SIMILARITY.
FT DISULFID 534 543 BY SIMILARITY.
FT DISULFID 550 560 BY SIMILARITY.
FT DISULFID 555 569 BY SIMILARITY.
FT DISULFID 571 580 BY SIMILARITY.
FT DISULFID 587 598 BY SIMILARITY.
FT DISULFID 592 607 BY SIMILARITY.
FT DISULFID 609 618 BY SIMILARITY.
FT DISULFID 625 635 BY SIMILARITY.
FT DISULFID 630 644 BY SIMILARITY.
FT DISULFID 646 655 BY SIMILARITY.
FT DISULFID 662 673 BY SIMILARITY.
FT DISULFID 667 682 BY SIMILARITY.
FT DISULFID 684 693 BY SIMILARITY.
FT DISULFID 700 710 BY SIMILARITY.
FT DISULFID 705 719 BY SIMILARITY.
FT DISULFID 721 730 BY SIMILARITY.
FT DISULFID 739 750 BY SIMILARITY.
FT DISULFID 744 759 BY SIMILARITY.
FT DISULFID 761 770 BY SIMILARITY.
FT DISULFID 776 787 BY SIMILARITY.
FT DISULFID 781 797 BY SIMILARITY.
FT DISULFID 799 808 BY SIMILARITY.
FT DISULFID 815 827 BY SIMILARITY.
FT DISULFID 821 836 BY SIMILARITY.

FT DISULFID 838 847 BY SIMILARITY.
FT DISULFID 854 865 BY SIMILARITY.
FT DISULFID 859 874 BY SIMILARITY.
FT DISULFID 876 885 BY SIMILARITY.
FT DISULFID 892 902 BY SIMILARITY.
FT DISULFID 897 911 BY SIMILARITY.
FT DISULFID 913 922 BY SIMILARITY.
FT DISULFID 929 940 BY SIMILARITY.
FT DISULFID 934 949 BY SIMILARITY.
FT DISULFID 951 960 BY SIMILARITY.
FT DISULFID 967 978 BY SIMILARITY.
FT DISULFID 972 987 BY SIMILARITY.
FT DISULFID 989 998 BY SIMILARITY.
FT DISULFID 1005 1016 BY SIMILARITY.
FT DISULFID 1010 1023 BY SIMILARITY.
FT DISULFID 1025 1034 BY SIMILARITY.
FT DISULFID 1041 1062 BY SIMILARITY.
FT DISULFID 1056 1071 BY SIMILARITY.
FT DISULFID 1073 1082 BY SIMILARITY.
FT DISULFID 1089 1100 BY SIMILARITY.
FT DISULFID 1094 1109 BY SIMILARITY.
FT DISULFID 1111 1120 BY SIMILARITY.
FT DISULFID 1127 1138 BY SIMILARITY.
FT DISULFID 1132 1147 BY SIMILARITY.
FT DISULFID 1149 1158 BY SIMILARITY.
FT DISULFID 1165 1183 BY SIMILARITY.
FT DISULFID 1177 1192 BY SIMILARITY.
FT DISULFID 1194 1203 BY SIMILARITY.

Query Match 39.1%; Score 84; DB 1; Length 2318;
Best Local Similarity 42.4%; Pred. No. 0.0099;
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGEVCIINNNHTCNCVDVGYPCQC 33
DB 929 CSPSSCFNGGCVGVSSFCICRGYGTGHQC 961

RESULT 33
ID PRTC_RAT STANDARD; PRT; 461 AA.
AC P31394;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE (Autoproteolysin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV).
GN PROC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=92329550; PubMed=1627650;
RA Okafuji T., Maekawa K., Nawa K., Marumoto Y.;
RT "The CDNA cloning and mRNA expression of rat protein C.";
CC Biochim. Biophys. Acta 1131:329-332(1992).
CC -i- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -i- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and VIIIA.
CC -i- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC STRONGLY PROMOTED BY THROMBOMODULIN.
CC -i- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -i- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME

CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
 CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
 CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
 CC SITE IS NECESSARY FOR THE RECOGNITION OF THE
 CC THROMBIN-THROMBOMODULIN COMPLEX.
 CC -1- SIMILARITY: CONSTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X64336; CAA45617.1; -.
 CC PIR: S18994; S18994.
 CC PIR: S24312; S24312.
 CC HSSP: P04070; LPCU.
 CC MEROPS: S01.218; -.
 CC InterPro: IPR000152; Asx_hydroxyl.
 CC InterPro: IPR001314; Chymotrypsin.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR001881; EGF_Ca.
 CC InterPro: IPR002383; GLA_blood.
 CC InterPro: IPR001254; Trypsin.
 CC InterPro: IPR000294; Vitk_dep_GLA.
 CC Pfam: PF00008; EGF; 2.
 CC Pfam: PF00594; gla; 1.
 CC Pfam: PF00089; trypsin; 1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC PRINTS: PR00001; GLABLOOD.
 CC SMART: SM00179; EGF_CA; 1.
 CC SMART: SM00001; EGF_Like; 1.
 CC SMART: SM00069; GLA; 1.
 CC SMART: SM00020; TRYP_SPE; 1.
 CC PROSITE: PS00010; ASX_HYDROXYL; 1.
 CC PROSITE: PS00022; EGF_1; 1.
 CC PROSITE: PS01186; EGF_2; 2.
 CC PROSITE: PS01187; EGF_CA; 1.
 CC PROSITE: PS00011; GLU_CARBOXYLATION; 1.
 CC PROSITE: PS00240; TRYPSIN_DOM; 1.
 CC PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC PROSITE: PS00135; TRYPSIN_SER; 1.
 CC Blood coagulation; Glycoprotein; Serine protease;
 CC Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
 CC EGF-like domain; Repeat; Endothelial cell; Hydrolyase; Signal.
 CC SIGNAL: 1 32
 CC PROPEP 33 41 BY SIMILARITY.
 CC CHAIN 42 196 PROTEIN C LIGHT CHAIN (BY SIMILARITY).
 CC CHAIN 199 461 PROTEIN C HEAVY CHAIN (BY SIMILARITY).
 CC PEPTIDE 199 212 ACTIVATION PEPTIDE (BY SIMILARITY).
 CC SITE 212 213 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 CC DOMAIN 96 131 EGF-LIKE 1.
 CC DOMAIN 135 175 EGF-LIKE 2.
 CC DOMAIN 213 461 SERINE PROTEASE.
 CC MOD_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID
 CC MOD_RES 48 48 GAMMA-CARBOXYGLUTAMIC ACID
 CC MOD_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID
 CC MOD_RES 57 57 GAMMA-CARBOXYGLUTAMIC ACID
 CC MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID
 CC MOD_RES 61 61 GAMMA-CARBOXYGLUTAMIC ACID
 CC MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID
 CC MOD_RES 67 67 GAMMA-CARBOXYGLUTAMIC ACID

FT MOD_RES 70 70 (BY SIMILARITY).
 FT ACT_SITE 112 112 GAMMA-CARBOXYGLUTAMIC ACID
 FT ACT_SITE 254 254 (BY SIMILARITY).
 FT ACT_SITE 300 300 HYDROXYLATION (BY SIMILARITY).
 FT ACT_SITE 402 402 CHARGE RELAY SYSTEM.
 FT DISULFID 58 63 CHARGE RELAY SYSTEM.
 FT DISULFID 91 110 BY SIMILARITY.
 FT DISULFID 100 105 BY SIMILARITY.
 FT DISULFID 104 119 BY SIMILARITY.
 FT DISULFID 121 130 BY SIMILARITY.
 FT DISULFID 139 150 BY SIMILARITY.
 FT DISULFID 146 159 BY SIMILARITY.
 FT DISULFID 161 174 BY SIMILARITY.
 FT DISULFID 182 320 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 239 255 BY SIMILARITY.
 FT DISULFID 373 387 BY SIMILARITY.
 FT DISULFID 398 426 BY SIMILARITY.
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 461 AA; 51912 MW; 8A4CF9364EDACD5 CRC64;
 Query Match 38.6%; Score 83; DB 1; Length 461;
 Best Local Similarity 42.9%; Pred No. 0.003; Indels 0; Gaps 0;
 Matches 12; Conservative 6; Mismatches 10;
 QY 6 CSGHGCEVEIINHNHCNDVGYGPOCQ 33
 DB 104 CCGHGTCTDGLGSGSCDKMGEGRCQ 131
 ID FBP3_STRPU STANDARD; PRT; 570 AA.
 AC P49013;
 DT 01-FEB-1996 (rel. 33, last sequence update)
 DT 16-OCT-2001 (rel. 40, last annotation update)
 DE Fibropellin C precursor (Epidermal growth factor-related protein 3)
 DE (EGF III) (Fibropellin III).
 GN EGF3.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinozoa;
 OC Echinoidea; Echinoidea; Echinacea; Echinacea; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=gastrula;
 RA MEDLINE=93273088; PubMed=8500658;
 RX Bisgrove B.W., Rafi R.A.;
 RT "The SPEGF III gene encodes a member of the fibropellins: EGF repeat-
 RT containing proteins that form the apical lamina of the sea urchin
 RT embryo.";
 RL Dev. Biol. 157:526-538(1993).
 CC -1- FUNCTION: FORM THE APICAL LAMINA, A COMPONENT OF THE EXTRACELLULAR
 CC MATRIX.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN UNFERTILIZED EGGS AND DURING
 CC EARLY CLEAVAGE, THEN RAPIDLY INCREASES IN ABUNDANCE BETWEEN LATE
 CC MORULA AND MESENCHYME BLASTULA STAGES TO MAXIMAL LEVELS MAINTAINED
 CC THROUGH SUBSEQUENT STAGES.
 CC -1- MISCELLANEOUS: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 CC -1- SIMILARITY: CONSTAINS 8 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONSTAINS 1 CUB DOMAIN.
 CC -1- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR
 CC TO AVIDIN/STREPTAVIDIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L07045; AAA30045.1; -.
 DR HSBP; P00740; IEDM.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000088; Avidin.
 DR InterPro: IPR000859; CUB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF-2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001438; EGF_II.
 DR Pfam; PF01382; Avidin; 1.
 DR Pfam; PF00431; CUB; 1.
 DR Pfam; PF00008; EGF; 8.
 DR PRINTS; PR00010; EGFBLDOD.
 DR ProDom; PD016055; Avidin; 1.
 DR SMART; SM00042; CUB; 1.
 DR SMART; SM00179; EGF_Ca; 7.
 DR SMART; SM00001; EGF-like; 1.
 DR PROSITE; PS00022; EGF_1; 8.
 DR PROSITE; PS00577; AVIDIN; 1.
 DR PROSITE; PS01180; CUB; 1.
 DR PROSITE; PS01186; EGF-2; 7.
 DR PROSITE; PS01187; EGF_Ca; 6.
 DR Biotin; EGF-like domain; Repeat; Signal; Glycoprotein;
 KW Calcium-binding.
 FT SIGNAL 1 17
 FT CHAIN 18 570
 FT DOMAIN 18 55
 FT DOMAIN 62 175
 FT DOMAIN 176 212
 FT DOMAIN 214 250
 FT DOMAIN 252 288
 FT DOMAIN 290 326
 FT DOMAIN 328 364
 FT DOMAIN 366 402
 FT DOMAIN 404 440
 FT DOMAIN 442 570
 FT DISULFID 23 34
 FT DISULFID 28 43
 FT DISULFID 45 54
 FT DISULFID 180 191
 FT DISULFID 185 200
 FT DISULFID 202 211
 FT DISULFID 218 229
 FT DISULFID 223 238
 FT DISULFID 240 249
 FT DISULFID 256 267
 FT DISULFID 261 276
 FT DISULFID 278 287
 FT DISULFID 294 305
 FT DISULFID 299 314
 FT DISULFID 316 325
 FT DISULFID 332 343
 FT DISULFID 337 352
 FT DISULFID 354 363
 FT DISULFID 370 381
 FT DISULFID 375 390
 FT DISULFID 392 401
 FT DISULFID 408 419
 FT DISULFID 413 428
 FT DISULFID 430 439
 FT CARBOHYD 30 30
 FT CARBOHYD 136 136
 FT CARBOHYD 357 357
 SQ SEQUENCE 570 AA: 61116 MW: BE653E3E1C05E6EE CRC64;

Query Match 38.6%; Score 83; DB 1; Length 570;
 Best Local Similarity 39.4%; Pred. No. 0.0036;
 Matches 13; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

OY 1 COPMSCSGHECVETIINHTCNCVGYGPOCQ 33

DB 218 CASAPCRNGACVDPVNGVTCNCIPFNGVNC 250

RESULT 35
 ID CRBH_HUMAN STANDARD; PRT; 1376 AA.
 AC P82279;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Crumbs protein homolog 1 precursor.
 GN CRBL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS RP12 V-161; W-250; M-745; C-764;
 RP Y-948; T-1041 AND P-1071.
 RC TISSUE=Retina, and Fetal brain;
 RX MEDLINE=99438399; PubMed=10508521;
 RA den Hollander A.I., ten Brink J.B., de Kok Y.J.M., van Soest S.,
 RA van den Born L.T., van Driel M.A., van de Pol D.J.R., Payne A.M.,
 RA Bhattacharya S.S., Kellner U., Hoyng C.B., Westerveld A.,
 RA Brunner H.G., Bleeker-Wagemakers E.M., Deutman A.F.,
 RA Heckenlied J.R., Cremers F.P.M., Bergen A.A.B.;
 RA "Mutations in a human homologue of drosophila crumbs cause retinitis
 RA pigmentosa (RP12)";
 RT Nat. Genet. 23:217-221(1999).
 RL [2]
 RN VARIANT LCA ARG-1100, AND VARIANT RP ARG-1181.
 RP MEDLINE=21303018; PubMed=11389483;
 RX den Hollander A.I., Heckenlied J.R., van den Born L.T.,
 RA de Kok Y.J.M., van der Velde-Visser S.D., Kellner U.,
 RA van Schooneveld M.J., Blankenagel A., Kohnscheider K., Wissing B.,
 RA Cruysberg J.R.M., Deutman A.F., Brunner H.G., Apfelstedt-Sylla E.,
 RA Hoyng C.B., Cremers F.P.M.;
 RA "Leber congenital amaurosis and retinitis pigmentosa with Coats-like
 RT exudative vasculopathy are associated with mutations in the crumbs
 RL homologue 1 (CRBL) gene";
 RL Am. J. Hum. Genet. 69:198-203(2001).
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTION IN NEURONAL
 CC DEVELOPMENT OF THE RETINA.
 CC -1- SUBCELLULAR LOCATION: Extracellular (Potential).
 CC -1- TISSUE SPECIFICITY: PREFERENTIAL EXPRESSION IN RETINA, ALSO
 CC EXPRESSED IN BRAIN AND FETAL BRAIN.
 CC -1- DISEASE: DEFECTS IN CRBL ARE THE CAUSE OF RETINITIS PIGMENTOSA
 CC TYPE 12 (RP12), AN AUTOSOMAL RECESSIVE CONDITION WHICH IS
 CC CHARACTERIZED BY NIGHT BLINDNESS FROM EARLY CHILDHOOD AND
 CC PROGRESSIVE VISUAL FIELD LOSS. THERE IS GENERAL LOSS OF RETINAL
 CC PIGMENT EPITHELIUM THROUGHOUT THE RETINA AND PATIENTS EXPERIENCE
 CC SEVERE VISUAL IMPAIRMENT BEFORE THE AGE OF TWENTY.
 CC -1- DISEASE: Defects in CRBL are a cause of a form of Leber congenital
 CC amaurosis (LCA).
 CC -1- SIMILARITY: CONTAINS 15 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
 CC -1- DATABASE: NAME-Mutations of the CRBL gene;
 CC NOTE=Retina International's Scientific Newsletter;
 CC WWW="http://www.retina-international.com/sci-news/crblmt.htm".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL: AF154671; AAF01361.1; -
DR HSSP: P08709; 1BP9.
DR MIM: 604210; -
DR MIM: 600105; -
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_U1.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR001791; Laminin_G.
DR Pfam: PF00008; EGF; 16.
DR Pfam: PF00054; Laminin_G; 3.
DR PRINTS: PR00010; EGFBLD.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00179; EGF_CA; 8.
DR SMART: SM00282; LamG; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 10.
DR PROSITE: PS00022; EGF_1; 15.
DR PROSITE: PS01186; EGF_2; 11.
DR PROSITE: PS01187; EGF_CA; 7.
DR PROSITE: PS50025; Lam_G_DOMAIN; 3.
KW EGF-like domain; Glycoprotein; Repeat; Signal; Disease mutation;
KW Retinitis pigmentosa; Vision.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1376 CRUMBS PROTEIN HOMOLOG 1.
FT DOMAIN 70 108 EGF-Like 1.
FT DOMAIN 110 146 EGF-Like 2.
FT DOMAIN 148 184 EGF-Like 3.
FT DOMAIN 186 222 EGF-Like 4.
FT DOMAIN 224 260 EGF-Like 5.
FT DOMAIN 304 337 EGF-Like 6.
FT DOMAIN 339 395 EGF-Like 7.
FT DOMAIN 441 481 EGF-Like 8.
FT DOMAIN 485 670 LAMININ G-Like 1.
FT DOMAIN 672 708 EGF-Like 9.
FT DOMAIN 714 885 LAMININ G-Like 2.
FT DOMAIN 887 923 EGF-Like 10.
FT DOMAIN 950 1137 LAMININ G-Like 3.
FT DOMAIN 1139 1175 EGF-Like 11.
FT DOMAIN 1177 1212 EGF-Like 12.
FT DOMAIN 1214 1250 EGF-Like 13.
FT DOMAIN 1255 1295 EGF-Like 14.
FT DOMAIN 1297 1333 EGF-Like 15.
FT DOMAIN 1333 1333 EGF-Like 15, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 74 85 POTENTIAL.
FT DISULFID 79 96 POTENTIAL.
FT DISULFID 98 107 POTENTIAL.
FT DISULFID 114 125 POTENTIAL.
FT DISULFID 119 134 POTENTIAL.
FT DISULFID 136 145 POTENTIAL.
FT DISULFID 152 163 POTENTIAL.
FT DISULFID 157 172 POTENTIAL.
FT DISULFID 174 183 POTENTIAL.
FT DISULFID 190 201 POTENTIAL.
FT DISULFID 195 210 POTENTIAL.
FT DISULFID 212 221 POTENTIAL.
FT DISULFID 228 239 POTENTIAL.
FT DISULFID 233 248 POTENTIAL.
FT DISULFID 250 259 POTENTIAL.
FT DISULFID 305 316 POTENTIAL.
FT DISULFID 310 325 POTENTIAL.
FT DISULFID 327 336 POTENTIAL.
FT DISULFID 343 354 POTENTIAL.
FT DISULFID 348 383 POTENTIAL.
FT DISULFID 385 394 POTENTIAL.
FT DISULFID 445 456 POTENTIAL.
FT DISULFID 450 469 POTENTIAL.
FT DISULFID 471 480 POTENTIAL.
FT DISULFID 676 687 POTENTIAL.
FT DISULFID 681 696 POTENTIAL.
FT DISULFID 698 707 POTENTIAL.

FT DISULFID 891 902 POTENTIAL.
FT DISULFID 896 911 POTENTIAL.
FT DISULFID 913 922 POTENTIAL.
FT DISULFID 1143 1154 POTENTIAL.
FT DISULFID 1148 1163 POTENTIAL.
FT DISULFID 1165 1174 POTENTIAL.
FT DISULFID 1181 1191 POTENTIAL.
FT DISULFID 1186 1200 POTENTIAL.
FT DISULFID 1202 1211 POTENTIAL.
FT DISULFID 1218 1229 POTENTIAL.
FT DISULFID 1223 1238 POTENTIAL.
FT DISULFID 1240 1249 POTENTIAL.
FT DISULFID 1259 1274 POTENTIAL.
FT DISULFID 1268 1283 POTENTIAL.
FT DISULFID 1285 1294 POTENTIAL.
FT DISULFID 1301 1312 POTENTIAL.
FT DISULFID 1306 1321 POTENTIAL.
FT DISULFID 1323 1332 POTENTIAL.
FT CARBOHYD 30 30 POTENTIAL.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 657 657 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 757 757 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 871 871 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 975 975 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1000 1000 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1190 1190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1265 1265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 A -> V (IN RP12).
FT VARIANT 250 250 C -> W (IN RP12).
FT VARIANT 745 745 /FTID-VAR_011642.
FT VARIANT 764 764 T -> M (IN RP12).
FT VARIANT 764 764 /FTID-VAR_011643.
FT VARIANT 948 948 R -> C (IN RP12).
FT VARIANT 948 948 C -> Y (IN RP12).
FT VARIANT 1041 1041 /FTID-VAR_011645.
FT VARIANT 1071 1071 M -> T (IN RP12).
FT VARIANT 1071 1071 /FTID-VAR_011646.
FT VARIANT 1100 1100 L -> P (IN RP12).
FT VARIANT 1100 1100 /FTID-VAR_011647.
FT VARIANT 1181 1181 I -> R (IN LCA).
FT VARIANT 1181 1181 /FTID-VAR_011648.
FT VARIANT 1181 1181 C -> R (IN RP WITH COAT-LIKE EXUDATIVE VASCULOPATHY).
FT SEQUENCE 1376 AA; 151412 MM; F380DF2AA046A2FE CRC64;
FT

Query Match 38.6%; Score 83; DB 1; Length 1376;
Best Local Similarity 39.4%; Pred. No. 0.0081;
Matches 13; Conservative 8; Mismatches 10; Indels 2; Gaps 1;

Qy 4 WS--CSGHECEVETINNHHCNCQDVGYYGQOOL 34

Db 306 WSKPCHNNATCEDSVNNTYCHWPGYTGACER 338

RESULT 36

CADN_DROME STANDARD: PRT; 3097 AA.

AC CADN_DROME; OSVJB7;
 AC 015943; 09VJB7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Neutral-cadherin precursor (Cadherin-N protein) (DN-cadherin).
 GN CAN OR CG7100.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Head, and Embryo;
 RX MEDLINE=97388431; PubMed=9247265;
 RA Iwai Y., Usui T., Hirano S., Steward R., Takeichi M., Uemura T.;
 RT "Axon patterning requires DN-cadherin, a novel neuronal adhesion
 receptor, in the Drosophila embryonic CNS.";
 RL Neuron 19:77-89(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Boltskov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Ciesley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paolo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
 RA Hosteln D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venier E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP INTERACTION WITH ARM.
 RX MEDLINE=98298928; PubMed=9635189;
 RA Loureiro J., Peifer M.;
 RT "Roles of Armadillo, a Drosophila catenin, during central nervous
 system development.";
 RL Curr. Biol. 8:622-632(1998).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE

CC SORTING OF HETEROGENEOUS CELL TYPES. MAY ASSOCIATE WITH ARM NEURAL
 CC ISOFORM AND PARTICIPATE IN THE TRANSMISSION OF DEVELOPMENTAL
 CC INFORMATION.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (potential).
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, THE PROTEIN FIRST APPEARS IN
 CC THE MESODERM AT STAGE 9 AND IS PRESENT IN THE PROBLASTS AND MUSCLE
 CC FIBERS BY STAGE 12 AND STAGE 14, RESPECTIVELY. AT STAGE 12 THE
 CC PROTEIN IS ALSO LOCATED IN THE AXONS OF THE ENTIRE CNS, BUT NOT IN
 CC THE GLIAL CELLS. IN THIRD INSTAR LARVAE PROTEIN IS EXPRESSED IN
 CC THE CNS NEUROPILE, PHOTORECEPTOR AXONS AND PRECURSORS OF ADULT
 CC MUSCLES.
 CC -1- SIMILARITY: CONTAINS 16 CADHERIN DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB002397; BAA22151.1; -;
 CC EMBL; AE003656; AAF53635.1; -;
 CC HSSP; P00740; IIXA.
 CC FLYBASE: FBgn0015609; Cadn.
 CC InterPro: IPR002126; Cadherin.
 CC InterPro: IPR000233; Cadherin_C-term.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR000742; EGF_2.
 CC InterPro: IPR001881; EGF_Ca.
 CC InterPro: IPR001791; Laminin_G.
 CC Pfam: PF00028; cadherin; 14.
 CC Pfam: PF01049; Cadherin_C-term; 1.
 CC Pfam: PF00008; EGF; 3.
 CC Pfam: PF00054; Laminin_G; 2.
 CC PRINTS; PR00205; CADHERIN.
 CC SMART; SM00112; CA; 16.
 CC SMART; SM00179; EGF_CA; 1.
 CC SMART; SM00001; EGF_Like; 3.
 CC SMART; SM00282; LamG; 2.
 CC PROSITE; PS00232; CADHERIN_1; 9.
 CC PROSITE; PS0268; CADHERIN_2; 16.
 CC PROSITE; PS00022; EGF_1; 3.
 CC PROSITE; PS0186; EGF_2; 3.
 CC PROSITE; PS50025; LamG_DOMAIN; 2.
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 CC Signal; EGF-like domain.
 CC SIGNAL; 1 36
 CC PROPEP 37 ?
 CC CHAIN ? 3097
 CC TRANSMEM 1454
 CC DOMAIN 1475
 CC DOMAIN 1476 3097
 CC DOMAIN 181 305
 CC DOMAIN 430 543
 CC DOMAIN 554 651
 CC DOMAIN 660 756
 CC DOMAIN 766 858
 CC DOMAIN 867 968
 CC DOMAIN 978 1078
 CC DOMAIN 1087 1183
 CC DOMAIN 1193 1299
 CC DOMAIN 1307 1414
 CC DOMAIN 1423 1514
 CC DOMAIN 1523 1630
 CC DOMAIN 1639 1742
 CC DOMAIN 1749 1861
 CC DOMAIN 1870 1966
 CC DOMAIN 1974 2085
 CC DOMAIN 2346 2377
 CC DOMAIN 2379 2585
 CC
 CC NEURAL-CADHERIN.
 CC EXTRACELLULAR (POTENTIAL).
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).
 CC CADHERIN 1.
 CC CADHERIN 2.
 CC CADHERIN 3.
 CC CADHERIN 4.
 CC CADHERIN 5.
 CC CADHERIN 6.
 CC CADHERIN 7.
 CC CADHERIN 8.
 CC CADHERIN 9.
 CC CADHERIN 10.
 CC CADHERIN 11.
 CC CADHERIN 12.
 CC CADHERIN 13.
 CC CADHERIN 14.
 CC CADHERIN 15.
 CC CADHERIN 16.
 CC EGF-Like 1.
 CC LAMININ G-Like 1.


```

FT DOMAIN 1346 1382 EGF-LIKE 35.
FT DOMAIN 1385 1423 EGF-LIKE 36.
FT REPEAT 1466 1486 LIN/NOTCH 1.
FT REPEAT 1487 1520 LIN/NOTCH 2.
FT REPEAT 1521 1561 LIN/NOTCH 3.
FT REPEAT 1667 1910 ANK 1.
FT REPEAT 1915 1944 ANK 2.
FT REPEAT 1948 1978 ANK 3.
FT REPEAT 1982 2011 ANK 4.
FT REPEAT 2015 2044 ANK 5.
FT REPEAT 2048 2077 ANK 6.
FT DOMAIN 2265 2276 POLY-GLN (OPA-REPEAT).
FT DISULFID 25 35 BY SIMILARITY.
FT DISULFID 29 45 BY SIMILARITY.
FT DISULFID 47 56 BY SIMILARITY.
FT DISULFID 62 73 BY SIMILARITY.
FT DISULFID 67 86 BY SIMILARITY.
FT DISULFID 88 97 BY SIMILARITY.
FT DISULFID 105 116 BY SIMILARITY.
FT DISULFID 110 126 BY SIMILARITY.
FT DISULFID 128 137 BY SIMILARITY.
FT DISULFID 143 154 BY SIMILARITY.
FT DISULFID 148 163 BY SIMILARITY.
FT DISULFID 165 174 BY SIMILARITY.
FT DISULFID 181 194 BY SIMILARITY.
FT DISULFID 188 203 BY SIMILARITY.
FT DISULFID 205 214 BY SIMILARITY.
FT DISULFID 221 232 BY SIMILARITY.
FT DISULFID 226 242 BY SIMILARITY.
FT DISULFID 244 253 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY.
FT DISULFID 265 280 BY SIMILARITY.
FT DISULFID 282 291 BY SIMILARITY.
FT DISULFID 298 311 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY.
FT DISULFID 322 331 BY SIMILARITY.
FT DISULFID 338 349 BY SIMILARITY.
FT DISULFID 343 358 BY SIMILARITY.
FT DISULFID 360 369 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY.
FT DISULFID 399 408 BY SIMILARITY.
FT DISULFID 415 428 BY SIMILARITY.
FT DISULFID 422 437 BY SIMILARITY.
FT DISULFID 439 448 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 460 475 BY SIMILARITY.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 503 BY SIMILARITY.
FT DISULFID 498 512 BY SIMILARITY.
FT DISULFID 514 523 BY SIMILARITY.
FT DISULFID 530 541 BY SIMILARITY.
FT DISULFID 535 550 BY SIMILARITY.
FT DISULFID 552 561 BY SIMILARITY.
FT DISULFID 568 578 BY SIMILARITY.
FT DISULFID 573 587 BY SIMILARITY.
FT DISULFID 589 598 BY SIMILARITY.
FT DISULFID 605 616 BY SIMILARITY.
FT DISULFID 610 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 653 BY SIMILARITY.
FT DISULFID 648 662 BY SIMILARITY.
FT DISULFID 664 673 BY SIMILARITY.
FT DISULFID 680 691 BY SIMILARITY.
FT DISULFID 685 700 BY SIMILARITY.
FT DISULFID 702 711 BY SIMILARITY.
FT DISULFID 716 728 BY SIMILARITY.
FT DISULFID 723 737 BY SIMILARITY.
FT DISULFID 739 748 BY SIMILARITY.
FT DISULFID 755 766 BY SIMILARITY.
FT DISULFID 760 775 BY SIMILARITY.
FT DISULFID 777 786 BY SIMILARITY.
FT DISULFID 793 804 BY SIMILARITY.

```

```

FT DISULFID 798 813 BY SIMILARITY.
FT DISULFID 815 824 BY SIMILARITY.
FT DISULFID 831 842 BY SIMILARITY.
FT DISULFID 836 853 BY SIMILARITY.
FT DISULFID 855 864 BY SIMILARITY.
FT DISULFID 871 882 BY SIMILARITY.
FT DISULFID 876 891 BY SIMILARITY.
FT DISULFID 893 902 BY SIMILARITY.
FT DISULFID 909 920 BY SIMILARITY.
FT DISULFID 914 929 BY SIMILARITY.
FT DISULFID 931 940 BY SIMILARITY.
FT DISULFID 947 958 BY SIMILARITY.
FT DISULFID 952 967 BY SIMILARITY.
FT DISULFID 969 978 BY SIMILARITY.
FT DISULFID 1023 1034 BY SIMILARITY.
FT DISULFID 1028 1043 BY SIMILARITY.
FT DISULFID 1045 1054 BY SIMILARITY.

Query Match 38.1%; Score 82; DB 1; Length 2437;
Best Local Similarity 32.4%; Pred. No. 0.018;
Matches 11; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

Oy 1 CQPMSCGHECEVETINNHTCMCDVGYGPQCOL 34
Db 947 CVSAPCRNGNCTDCVNSYTCSCPAGFGINCET 980

RESULT 38
FA12_HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (Haf).
GN F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGilivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region."
RT J. Biol. Chem. 262:13662-13673(1987).
RN [2]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RT Nucleic Acids Res. 14:3146-3146(1986).
RN [3]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGilivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor Xlla."
RT J. Biol. Chem. 260:13666-13676(1985).
RN [4]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RT Biochemistry 25:1525-1528(1986).
RN [5]

```

RP SEQUENCE OF 20-379.
RX MEDLINE-85182674; PubMed-3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor).";
RL J. Biol. Chem. 260:5328-5341(1985).
RN [6]
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE-83291041; PubMed-6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa.";
RL J. Biol. Chem. 258:10924-10933(1983).
RN [7]
RP SEQUENCE OF 561-615 FROM N.A.
RC TISSUE-Blood;
RX MEDLINE-96133302; PubMed-8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G-->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material
RT negative patients.";
RL Hum. Mol. Genet. 4:1235-1237(1995).
RN [8]
RP VARIANT WASHINGTON DC.
RX MEDLINE-90046788; PubMed-2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alying B.,
RA Saito H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT factor XIIa results from Cys-571-->Ser substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
RN [9]
RP VARIANT LOCARNO.
RX MEDLINE-94325559; PubMed-8049433;
RA Hovington J.K., Schaller J., Stricker H., Wullemmin W.A., Furian M.,
RA Laemmle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by
RT the amino acid substitution Arg-353-->Pro leading to loss of a
RT kallikrein cleavage site.";
RL Blood 84:1173-1181(1994).
RN [10]
RP CARBOHYDRATE-LINKAGE SITE THR-109.
RX MEDLINE-92184750; PubMed-1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [11]
RP FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC [12]
CC CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC [13]
CC PTM: O- AND N-GLYCOSYLATED.
CC [14]
CC DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
CC SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC [15]
CC MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIa AND THEN TO BETA-FACTOR XIIa. ALPHA-FACTOR
CC XIIa ACTIVATES FACTOR XI TO FACTOR XIa.
CC [16]
CC SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC [17]
CC SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC [18]
CC SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC [19]
CC SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC [20]
CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC [21]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: M31315; AAA70225.1; -;
DR EMBL: M11723; AAA51986.1; -;
DR EMBL: M17466; AAB59490.1; -;
DR EMBL: M17464; AAB59490.1; JOINED.
DR EMBL: M17465; AAB59490.1; JOINED.
DR EMBL: M13147; AAA70224.1; -;
DR EMBL: U71274; AAB51203.1; -;
DR PIR: A29411; KFH12.
DR HSSP: P00750; 1RTE.
DR MEROPS: S01.211; -;
DR MIM: 234000; -;
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001294; Trypsin.
DR InterPro: IPR000083; Fibrinectin_type_1.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00013; ENTPEIT.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD00995; FN_Type_II; 1.
DR SMART: SM00181; EGF_2.
DR SMART: SM00058; FN1; 1.
DR SMART: SM00059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00023; FIBRONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HTS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR GlycoProtex: Blood coagulation; Plasma; Kringle; Serine protease;
DR HydroLase: Fibrinolysis; Signal; EGF-like domain; Repeat; zymogen;
DR Disease mutation.
DR SIGNAL 1 19
FT CHAIN 20 372 ALPHA-FACTOR XIIa HEAVY CHAIN.
FT CHAIN 373 615 ALPHA-FACTOR XIIa LIGHT CHAIN.
FT CHAIN 354 362 BETA-FACTOR XIIa PART 1.
FT CHAIN 373 615 BETA-FACTOR XIIa PART 2.
FT DOMAIN 47 88 FIBRONECTIN TYPE-I.
FT DOMAIN 94 131 EGF-LIKE 1.
FT DOMAIN 133 173 FIBRONECTIN TYPE-II.
FT DOMAIN 174 210 EGF-LIKE 2.
FT DOMAIN 217 295 KRINGLE.
FT DOMAIN 296 349 PRO-RICH.
FT DOMAIN 373 615 SERINE PROTEASE.
FT CARBOHYD 109 109 N-LINKED (RUC).
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .).
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 563 563 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 98 110 BY SIMILARITY.
FT DISULFID 104 119 BY SIMILARITY.
FT DISULFID 121 130 BY SIMILARITY.
FT DISULFID 135 163 BY SIMILARITY.
FT DISULFID 161 170 BY SIMILARITY.
FT DISULFID 178 189 BY SIMILARITY.
FT DISULFID 183 198 BY SIMILARITY.

ID NTCL_HUMAN STANDARD; PRT; 2444 AA.
AC P46531;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neogenenic locus notch protein homolog 1 precursor (Translocation-
GN associated notch protein TAN-1) (Fragment).
DE NOTCH1 OR TAN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91347367; PubMed=1831692;
RA Ellisen L.W., Bird J., West D.C., Soreng A.L., Reynolds T.C.,
RA Smolch S.D., Sklar J.;
RT "TAN-1, the human homolog of the Drosophila notch gene, is broken by
RT chromosomal translocations in T lymphoblastic neoplasms.";
RL Cell 66:649-661(1991).
CC -1- FUNCTION: MAY BE IMPORTANT FOR NORMAL LYMPHOCYTE FUNCTION. IN
CC ALTERED FORM, MAY CONTRIBUTE TO TRANSFORMATION OR PROGRESSION
CC IN SOME T-CELL NEOPLASMS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: IN FETAL TISSUES MOST ABUNDANT IN SPLEEN,
CC BRAIN STEM AND LUNG. ALSO PRESENT IN MOST ADULT TISSUES WHERE IT
CC IS FOUND MAINLY IN LYMPHOID TISSUES.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL; M73980; AAA0614.1; -.
DR HSSP; P00740; IIXA.
DR MIM; 190198; -.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000800; Notch.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 36.
DR Pfam; PF00066; notch; 3.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF_CA; 22.
DR SMART; SM00001; EGF_Like; 13.
DR SMART; SM00004; NL; 2.
DR PROSITE; PS50086; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 20.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 26.
DR PROSITE; PS01187; EGF_CA; 18.
DR Differentiation; Neogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 >2444 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 1.
FT DOMAIN 19 1736 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1737 1757 POTENTIAL.
FT DOMAIN 1758 >2444 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 58 EGF-LIKE 1.
FT DOMAIN 59 99 EGF-LIKE 2.
FT DOMAIN 102 139 EGF-LIKE 3.
FT DOMAIN 140 176 EGF-LIKE 4.

FT DOMAIN 178 216 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 218 255 EGF-LIKE 6.
FT DOMAIN 257 293 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 295 333 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 335 371 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 372 410 EGF-LIKE 10.
FT DOMAIN 412 450 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 452 488 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 490 526 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 528 564 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 566 601 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 603 639 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 641 676 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 678 714 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 716 751 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 753 789 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 791 827 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 829 868 EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 870 906 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 908 944 EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 946 982 EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 984 1020 EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1022 1058 EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1060 1096 EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1098 1144 EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1146 1182 EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1184 1220 EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1222 1266 EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1268 1306 EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1308 1347 EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1349 1385 EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1388 1427 EGF-LIKE 36, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1446 1481 LIN/NOTCH 1.
FT REPEAT 1482 1523 LIN/NOTCH 2.
FT REPEAT 1524 1563 LIN/NOTCH 3.
FT REPEAT 1565 1599 ANK 1.
FT REPEAT 1597 1633 ANK 2.
FT REPEAT 1635 1669 ANK 3.
FT REPEAT 1671 1705 ANK 4.
FT REPEAT 1707 1741 ANK 5.
FT DOMAIN 1743 1779 POLY-VAL.
FT DOMAIN 1781 1817 POLY-ARG.
FT DOMAIN 1819 1855 POLY-PRO.
FT DOMAIN 1857 1893 POLY-ALA.
FT DOMAIN 1895 1931 POLY-GLU.
FT DOMAIN 1933 1969 POLY-GLY.
FT DOMAIN 1971 2007 POLY-GLN.
FT DOMAIN 2009 2045 POLY-PRO.
FT DISULFID 2047 2083 BY SIMILARITY.
FT DISULFID 2085 2121 BY SIMILARITY.
FT DISULFID 2123 2159 BY SIMILARITY.
FT DISULFID 2161 2197 BY SIMILARITY.
FT DISULFID 2199 2235 BY SIMILARITY.
FT DISULFID 2237 2273 BY SIMILARITY.
FT DISULFID 2275 2311 BY SIMILARITY.
FT DISULFID 2313 2349 BY SIMILARITY.
FT DISULFID 2351 2387 BY SIMILARITY.
FT DISULFID 2389 2425 BY SIMILARITY.
FT DISULFID 2427 2463 BY SIMILARITY.
FT DISULFID 2465 2501 BY SIMILARITY.
FT DISULFID 2503 2539 BY SIMILARITY.
FT DISULFID 2541 2577 BY SIMILARITY.
FT DISULFID 2579 2615 BY SIMILARITY.
FT DISULFID 2617 2653 BY SIMILARITY.
FT DISULFID 2655 2691 BY SIMILARITY.
FT DISULFID 2693 2729 BY SIMILARITY.
FT DISULFID 2731 2767 BY SIMILARITY.
FT DISULFID 2769 2805 BY SIMILARITY.
FT DISULFID 2807 2843 BY SIMILARITY.
FT DISULFID 2845 2881 BY SIMILARITY.
FT DISULFID 2883 2919 BY SIMILARITY.
FT DISULFID 2921 2957 BY SIMILARITY.
FT DISULFID 2959 2995 BY SIMILARITY.
FT DISULFID 2997 3033 BY SIMILARITY.
FT DISULFID 3035 3071 BY SIMILARITY.
FT DISULFID 3073 3109 BY SIMILARITY.
FT DISULFID 3111 3147 BY SIMILARITY.
FT DISULFID 3149 3185 BY SIMILARITY.
FT DISULFID 3187 3223 BY SIMILARITY.
FT DISULFID 3225 3261 BY SIMILARITY.
FT DISULFID 3263 3299 BY SIMILARITY.
FT DISULFID 3301 3337 BY SIMILARITY.
FT DISULFID 3339 3375 BY SIMILARITY.

FT	DISULEID	344	359	BY SIMILARITY.
FT	DISULEID	361	370	BY SIMILARITY.
FT	DISULEID	376	387	BY SIMILARITY.
FT	DISULEID	381	398	BY SIMILARITY.
FT	DISULEID	400	409	BY SIMILARITY.
FT	DISULEID	416	429	BY SIMILARITY.
FT	DISULEID	423	438	BY SIMILARITY.
FT	DISULEID	440	449	BY SIMILARITY.
FT	DISULEID	456	467	BY SIMILARITY.
FT	DISULEID	461	476	BY SIMILARITY.
FT	DISULEID	478	487	BY SIMILARITY.
FT	DISULEID	494	505	BY SIMILARITY.
FT	DISULEID	499	514	BY SIMILARITY.
FT	DISULEID	516	525	BY SIMILARITY.
FT	DISULEID	532	543	BY SIMILARITY.
FT	DISULEID	537	552	BY SIMILARITY.
FT	DISULEID	554	563	BY SIMILARITY.
FT	DISULEID	570	580	BY SIMILARITY.
FT	DISULEID	575	589	BY SIMILARITY.
FT	DISULEID	591	600	BY SIMILARITY.
FT	DISULEID	607	618	BY SIMILARITY.
FT	DISULEID	612	627	BY SIMILARITY.
FT	DISULEID	629	638	BY SIMILARITY.
FT	DISULEID	645	655	BY SIMILARITY.
FT	DISULEID	650	664	BY SIMILARITY.
FT	DISULEID	666	675	BY SIMILARITY.
FT	DISULEID	682	693	BY SIMILARITY.
FT	DISULEID	687	702	BY SIMILARITY.
FT	DISULEID	704	713	BY SIMILARITY.
FT	DISULEID	720	730	BY SIMILARITY.
FT	DISULEID	725	739	BY SIMILARITY.
FT	DISULEID	741	750	BY SIMILARITY.
FT	DISULEID	757	768	BY SIMILARITY.
FT	DISULEID	762	777	BY SIMILARITY.
FT	DISULEID	779	788	BY SIMILARITY.
FT	DISULEID	795	806	BY SIMILARITY.
FT	DISULEID	800	815	BY SIMILARITY.
FT	DISULEID	817	826	BY SIMILARITY.
FT	DISULEID	833	844	BY SIMILARITY.
FT	DISULEID	838	855	BY SIMILARITY.
FT	DISULEID	857	867	BY SIMILARITY.
FT	DISULEID	874	885	BY SIMILARITY.
FT	DISULEID	879	894	BY SIMILARITY.
FT	DISULEID	896	905	BY SIMILARITY.
FT	DISULEID	912	923	BY SIMILARITY.
FT	DISULEID	917	932	BY SIMILARITY.
FT	DISULEID	934	943	BY SIMILARITY.
FT	DISULEID	988	999	BY SIMILARITY.
FT	DISULEID	993	1008	BY SIMILARITY.
FT	DISULEID	1010	1019	BY SIMILARITY.
FT	DISULEID	1026	1037	BY SIMILARITY.
FT	DISULEID	1031	1046	BY SIMILARITY.
FT	DISULEID	1048	1057	BY SIMILARITY.
FT	DISULEID	1064	1075	BY SIMILARITY.

Query Match 37.7%; Score 81; DB 1; Length 2444;
 Best Local Similarity 45.5%; Pred. No. 0.024;
 Matches 15; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 CQWSCSGHCEVETINNHTCNCVGYGPOCQ 33
 Db 988 CTSSCFMGTCVDGINSFTCLCPGFGSYCQ 1020

Search completed: September 7, 2002, 10:23:43
 Job time: 536 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:15:25 : Search time 36.49 Seconds
(without alignments)
89.532 Million cell updates/sec

Title: US-09-119-209-2_COPY_160_193

Perfect score: 215

Sequence: 1 COPWSCSGHGECVEIINNHTCNCDDVGYGPOCOL 34

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR_71:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	205	95.3	323	1	S09702 L-selectin precurs
2	205	95.3	372	2	JC5377 L-selectin precurs
3	205	95.3	385	1	A34015 L-selectin precurs
4	191	88.8	370	2	S22124 L-selectin precurs
5	181	84.2	376	2	JC4892 L-selectin precurs
6	169	78.6	372	1	A32375 L-selectin precurs
7	159	74.0	372	2	S23936 L-selectin precurs
8	130	60.5	610	2	A35046 E-selectin precurs
9	127	59.1	646	2	JN0473 E-selectin precurs
10	126	58.6	612	2	B42755 E-selectin precurs
11	125	58.1	482	2	JC5092 E-selectin precurs
12	125	58.1	830	2	A30359 E-selectin - pig
13	124	57.7	768	2	A42755 P-selectin precurs
14	118	54.9	485	2	S36772 E-selectin - bovin
15	116	54.0	551	2	I46709 endothelial leukoc
16	109	50.7	768	2	I53821 P-selectin - rat
17	95	44.2	603	2	S28941 coagulation factor
18	95	44.2	710	2	T21339 hypothetical prote
19	93	43.3	2524	2	A35844 Notch protein - Af
20	90	41.9	2531	2	T31070 notch homolog - se
21	89	41.9	3051	2	S42373 hypothetical prote
22	89	41.4	1203	2	A49175 Notch B protein -
23	89	41.4	2321	2	S78549 notch3 protein - h
24	89	41.4	2471	2	A49128 cell-fate determin
25	87	40.5	461	1	JN0210 protein C (activat
26	87	40.5	728	2	I30719 C-beta-1 - chicke
27	87	40.5	1064	2	A40136 fibropellin Ia - s
28	86	40.0	2139	2	A35672 crumbs protein - f
29	86	40.0	2703	1	A24420 notch protein - fr

30	85.5	39.8	473	2	A56175
31	85.5	39.8	3461	2	S58870
32	85	39.5	1722	2	E89753
33	85	39.5	2352	2	T30201
34	84	39.1	372	2	T29359
35	84	39.1	2019	1	JQ1322
36	84	39.1	2201	2	A32160
37	84	39.1	2318	2	S45306
38	84	39.1	3623	2	T09456
39	83.5	38.8	1069	2	T42681
40	83.5	38.8	1523	2	T13953
41	83.5	38.8	2809	2	T30213
42	83	38.6	293	2	B26337
43	83	38.6	461	1	S18994
44	83	38.6	570	2	A48836
45	83	38.6	3097	2	T00021

ALIGNMENTS

RESULT 1
S09702 L-selectin precursor, short splice form - human
N/Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukoc
ral lymph node homing receptor Leu-8
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 26-May-2000
R/Camerini, D.; James, S.P.; Stamenkovic, I.; Seed, B.
Nature 342, 78-82, 1989
A/Title: Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor.
A/Reference number: S06796; MUID:90044046
A/Accession: S09702
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-323 <CAM>
A/Cross-references: EMBL:X17519; NID:g34344
A/Note: this translation is not annotated in GenBank entry HSLEU8, release 111.0
C/Comment: For an alternative splice form, see PIR:A34015.
C/Genetics:
A/Gene: GDB:SEL; GDB:LNHR; LSEL; LAM1; LYAM1; LAM-1
A/Cross-references: GDB:120157; GDB:118834; OMIM:153240
A/Map position: 1q22-1q23
A/Intons: 14/3; 42/1; 171/1; 207/1; 269/1; 303/2
C/Function:
A/Description: binds with low affinity to oligosaccharides like heparan sulfate and si
ment of leucocytes to areas of inflammation, and with CD162 mediates neutrophil-neutr
C/Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology
C/Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammat
F/1-51/Domain: signal sequence #status predicted <SIG>
F/42-168/Domain: C-type lectin homology <LCH>
F/52-323/Product: L-selectin #status predicted <MAT>
F/52-300/Domain: extracellular #status predicted <EXT>
F/173-204/Domain: EGF homology <EGF>
F/210-267/Domain: complement factor H repeat homology <FH1>
F/272-323/Domain: complement factor H repeat homology #status atypical <FH2>
F/301-318/Domain: transmembrane #status predicted <TRM>
F/319-323/Domain: intracellular #status predicted <INT>
F/73-117,190,245,259/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/322/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 95.3%; Score 205; DB 1; Length 323;
Best Local Similarity 97.0%; Pred. No. 7e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECVEIINNHTCNCDDVGYGPOCOL 33
DB 173 COPWSCSGHGECVEIINNHTCNCDDVGYGPOCOL 205
RESULT 2

JC5377
L-selectin precursor - hamadryas baboon
C:Species: Papio hamadryas (hamadryas baboon)
C>Date: 02-Jun-1997 #sequence_revision 12-Sep-1997 #text_change 21-Jan-2000
C:Accession: JC5377; PC4315
R:tsurushita, N.; Fu, H.; Berg, E.L.
Gene 181, 219-220, 1996
A:Title: PCR cloning of the cDNA encoding baboon L-selectin.
A:Reference number: JC5377; MUID:97128794
A:Accession: JC5377
A:Molecule type: mRNA
A:Residues: 1-372 <TSU1>
A:Cross-references: GB:052074; NID:g1326148; PIDN:AAB40903.1; PID:g1326149
A:Accession: PC4315
A:Molecule type: protein
A:Residues: 37-43:142-148 <TSU2>
C:Comment: This receptor is involved in the initial adhesive interaction between lymphoc sites of inflammation.
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;
F:1-38/Domain: signal sequence #status predicted <SIG>
F:29-155/Domain: C-type lectin homology <LCH>
F:39-372/Product: L-selectin #status predicted <MAT>
F:39-157/Domain: calcium-binding #status predicted <CAB>
F:160-191/Domain: EGF homology <EGF>
F:197-254/Domain: complement factor H repeat homology <FH1>
F:259-316/Domain: complement factor H repeat homology <FH2>
F:333-355/Domain: transmembrane #status predicted <TM>
F:356-372/Domain: intracellular #status predicted <INT>

Query Match 95.3%; Score 205; DB 2; Length 372;
Best Local Similarity 97.0%; Pred. No. 7.9e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQWSCSGHGEVCYEIIINNHCNCDVGYGPQCQ 33
Db 160 CQWSCSGHGEVCYEIIINNHCNCDVGYGPQCQ 192

RESULT 3

A34015
L-selectin precursor, long splice form - human
N:Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukocyte
ral lymph node homing receptor Leu-8
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: I55333; S06798; JI0104; A34015; A33912
R:Ord, D.C.; Ernst, T.J.; Zhou, L.J.; Rambaldi, A.; Spertini, O.; Griffin, J.; Tedder, T.
J. Biol. Chem. 265, 7760-7767, 1990
A:Title: Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Le
A:Reference number: I55333; MUID:90243637
A:Accession: I55333
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 14-385 <ORD>
A:Cross-references: GB:I32414; NID:g187259; PIDN:AAB60700.1; PID:g386860
R:Cammerlin, D.; James, S.P.; Stamenkovic, I.; Seed, B.
Nature 342, 78-82, 1989
A:Title: Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor.
A:Reference number: S06798; MUID:90044046
A:Accession: S06798

A:Molecule type: mRNA
A:Residues: 1-225, 'S', 227-385 <CAM>
A:Cross-references: EMBL:X17519; NID:g34344; PIDN:CAB43536.1; PID:g4902829
A:Note: this translation is not annotated in GenBank entry HSLER8, release 111.0
R:tedder, T.F.; Isaacs, C.M.; Ernst, T.J.; Demetri, G.D.; Adler, D.A.; Distche, C.M.
J. Exp. Med. 170, 123-133, 1989
A:Title: Isolation and chromosomal localization of cDNAs encoding a novel human lymphocy
oteins.
A:Reference number: JI0104; MUID:89310350
A:Accession: JI0104
A:Molecule type: mRNA
A:Residues: 1-230, 'N', 232, 'N', 234-254, 'E', 256-385 <TED>

A:Cross-references: GB:X16150; NID:g34428; PIDN:CAA34275.1; PID:g34429
A:Note: the translated sequence in GenBank entry HSLYML, release 111.0, differs from
R:Bowen, B.R.; Nguyen, T.; Lasky, L.A.
J. Cell Biol. 109, 421-427, 1989
A:Title: Characterization of a human homologue of the murine peripheral lymph node ho
A:Reference number: A34015; MUID:89308881
A:Accession: A34015
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 14-49, 'V', 51-190, 'H', 192-205, 'L', 207-226, 'E', 228-385 <ROW>
A:Cross-references: GB:X16070; NID:g38092; PIDN:CAA34203.1; PID:g38093
R:Siegelman, M.H.; Weissman, I.L.
Proc. Natl. Acad. Sci. U.S.A. 86, 5562-5566, 1989
A:Title: Human homologue of mouse lymph node homing receptor: evolutionary conservati
A:Reference number: A33912; MUID:89315837
A:Accession: A33912
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 14-205, 'L', 207-385 <SIE>
A:Cross-references: GB:M25280; NID:g187182; PIDN:AAC63053.1; PID:g307134
C:Comment: For an alternative splice form, see PIR:S09702.
C:Genetics:
A:Gene: GDB:SELL; GDB:LNHR; ISFL; LAM1; LYAM1; LAM-1
A:Cross-references: GDB:120157; GDB:118834; OMIM:153240
A:Map position: 1922-1923
A:Intons: 14/3; 42/1; 171/1; 207/1; 269/1; 331/1; 374/1; 380/2
C:Function:
A:Description: binds with low affinity to oligosaccharides like heparan sulfate and si
ment of leukocytes to areas of inflammation, and with CD162 mediates neutrophil-neutr
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homolog
C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammat
F:1-51/Domain: signal sequence #status predicted <SIG>
F:42-168/Domain: C-type lectin homology <LCH>
F:52-385/Product: L-selectin #status predicted <MAT>
F:52-343/Domain: extracellular #status predicted <EXT>
F:173-204/Domain: EGF homology <EGF>
F:210-267/Domain: complement factor H repeat homology <FH1>
F:272-329/Domain: complement factor H repeat homology <FH2>
F:344-368/Domain: transmembrane #status predicted <TM>
F:369-385/Domain: intracellular #status predicted <INT>
F:73, 117, 190, 245, 259/Binding site: carbohydrate (asn) (covalent) #status predicted
F:377, 380/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 95.3%; Score 205; DB 1; Length 385;
Best Local Similarity 97.0%; Pred. No. 8.1e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQWSCSGHGEVCYEIIINNHCNCDVGYGPQCQ 33
Db 173 CQWSCSGHGEVCYEIIINNHCNCDVGYGPQCQ 205

RESULT 4

S22124
L-selectin precursor - bovine
N:Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)
C:Species: Bos primigenius taurus (cattle)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jan-2000
C:Accession: S22124; A46531
R:Bosworth, B.T.
submitted to the EMBL Data Library, October 1991
A:Reference number: S22123
A:Accession: S22124
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-370 <BOS>
A:Cross-references: EMBL:X62882; NID:g515; PIDN:CAA44676.1; PID:g516
R:Witcher, B.; White, M.; Kurk, S.; Kishimoto, T.K.; Jutila, M.A.
Eur. J. Immunol. 22, 469-476, 1992
A:Title: Characterization of the bovine peripheral lymph node homing receptor: a lect
A:Reference number: A46531; MUID:92164727
A:Accession: A46531

Query Match 74.0%; Score 159; DB 2; Length 372;
Best Local Similarity 78.8%; Pred. No. 1.1e-10;
Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVCETIINNNHTCNCVGYGPQCQ 33
DB 160 CQPMSCNRHCEVCETIINNNHTCNCVGYGPQCQ 192

RESULT 8

A35046
E-selectin precursor - human
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1); LECAM-2
C:Species: Homo sapiens (man)
C>Date: 27-Jul-1990 #sequence, revision 27-Jul-1990 #text_change 21-Jul-2000
C:Accession: A38615; A35046; A32606
R:Collins, T.; Williams, A.; Johnston, G.I.; Kim, J.; Eddy, R.; Shows, T.; Gimbrone Jr., J. Biol. Chem. 266, 2466-2473, 1991
A:Title: Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule
A:Reference number: A38615; MUID:91115870
A:Accession: A38615
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-610 <COL>
A:Cross-references: GB:M61893; GB:M58017; NID:9182043; PIDN:AAA52375.1; PID:9182046
R:Hession, C.; Osborn, L.; Goff, D.; Chi-Rosso, G.; Vassallo, C.; Pasek, M.; Pittack, C. Proc. Natl. Acad. Sci. U.S.A. 87, 1673-1677, 1990
A:Title: Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional characterization
A:Reference number: A35046; MUID:90175359
A:Accession: A35046
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-610 <HES>
A:Cross-references: GB:M30640; NID:9182047; PIDN:AAA52377.1; PID:9182048
R:Bevilacqua, M.P.; Stangelin, S.; Gimbrone Jr., M.A.; Seed, B. Science 243, 1160-1165, 1989
A:Title: Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils
A:Reference number: A32606; MUID:89162047
A:Accession: A32606
A:Molecule type: mRNA
A:Residues: 1-467, 'Y', 469-610 <BEV>
A:Cross-references: GB:M24736; NID:9537523; PIDN:AAA52376.1; PID:9537524
C:Genetics:
A:Gene: GDB:SELE; ELAM; ESEL; ELAM1
A:Cross-references: GDB:120612; OMIM:131210
A:Map position: 1q22-1q25
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor
C:Keywords: duplication; glycoprotein; tandem repeat; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:12-138/Domain: C-type lectin homology <LCH>
F:22-610/Product: endothelial leukocyte adhesion molecule 1 #status predicted <MAT>
F:143-174/Domain: EGF homology <EGF>
F:180-237/Domain: complement factor H repeat homology <FH01>
F:242-299/Domain: complement factor H repeat homology <FH02>
F:304-362/Domain: complement factor H repeat homology <FH03>
F:367-425/Domain: complement factor H repeat homology <FH04>
F:430-488/Domain: complement factor H repeat homology <FH05>
F:493-547/Domain: complement factor H repeat homology <FH06>
F:557-578/Domain: transmembrane #status predicted <TM>
F:25,145,160,179,199,203,265,312,332,503,527/Binding site: carbohydrate (Asn) (covalent)

Query Match 60.5%; Score 130; DB 2; Length 610;
Best Local Similarity 63.6%; Pred. No. 3e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVCETIINNNHTCNCVGYGPQCQ 33
DB 143 CQPMSCSGHCEVCETIINNNHTCNCVGYGPQCQ 175

RESULT 9

JN0473
P-selectin precursor - bovine
N:Alternate names: granule membrane protein-140
C:Species: Bos primigenius taurus (cattle)
C>Date: 30-Sep-1993 #sequence, revision 30-Sep-1993 #text_change 11-Jan-2000
C:Accession: JN0473
R:Strubel, N.A.; Nguyen, M.; Kansas, G.S.; Tedder, T.F.; Bischoff, J. Biochem. Biophys. Res. Commun. 192, 338-344, 1993
A:Title: Isolation and characterization of a bovine cDNA encoding a functional homologue of human P-selectin
A:Reference number: JN0473; MUID:93249394
A:Accession: JN0473
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-646 <STR>
A:Cross-references: GB:L12041; NID:9304246; PIDN:AAA30743.1; PID:9304247
C:Comment: This protein is a Ca2+ dependent receptor for myeloid cells.
C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;
C:Keywords: cell adhesion; glycoprotein; phosphatidyl; phosphoprotein; transmembrane
F:1-41/Domain: signal sequence #status predicted <SIG>
F:42-646/Product: P-selectin #status predicted <MAT>
F:163-194/Domain: EGF homology <EGF>
F:200-257/Domain: complement factor H repeat homology <FH1>
F:262-319/Domain: complement factor H repeat homology <FH2>
F:324-381/Domain: complement factor H repeat homology <FH3>
F:386-443/Domain: complement factor H repeat homology <FH4>
F:458-515/Domain: complement factor H repeat homology <FH5>
F:520-577/Domain: complement factor H repeat homology <FH6>
F:588-611/Domain: transmembrane #status predicted <TM>
F:612-646/Domain: intracellular #status predicted <CYT>

Query Match 59.1%; Score 127; DB 2; Length 646;
Best Local Similarity 57.6%; Pred. No. 6.9e-07;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVCETIINNNHTCNCVGYGPQCQ 33
DB 163 CQPMSCSGHCEVCETIINNNHTCNCVGYGPQCQ 195

RESULT 10

B42755
E-selectin precursor - mouse
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)
C:Species: Mus musculus (house mouse)
C>Date: 04-Mar-1993 #sequence, revision 31-Dec-2000 #text_change 31-Dec-2000
C:Accession: S23174; B42755
R:Becker Andre, M.; van Huijsduijnen, R.H.; Losberger, C.; Whelan, J.; Delamarier, J. Eur. J. Biochem. 206, 401-411, 1992
A:Title: Murine endothelial leukocyte-adhesion molecule 1 is a close structural and functional homologue of human E-selectin
A:Reference number: S23174; MUID:92283265
A:Accession: S23174
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <BEC>
A:Cross-references: GB:M80778; NID:9193014; PIDN:AAA37547.1; PID:9193015
R:Weller, A.; Isenmann, S.; Vestweber, D. J. Biol. Chem. 267, 15176-15183, 1992
A:Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin in endothelial cells
A:Reference number: A42755; MUID:92340571
A:Accession: B42755
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'MKATAGV', 1-389, 391-612 <WEL>
A:Cross-references: GB:M87862; NID:9193107
A:Experimental source: endothelial cells
A:Note: Sequence extracted from NCBI backbone (NCBI:109470)
A:Note: the sequence in GenBank entry MUSELELC, release 117.0, (PIDN:AAA37577.1; PID:9193015) is uncertain whether the initiator is Met-1 or the AUG codon preceding the initiator
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor
C:Keywords: glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:12-138/Domain: C-type lectin homology <LCH>

F.22-612/Product: P-selectin #status predicted <MAT>
 F.143-174/Domain: EGF homology <EGF>
 F.180-238/Domain: complement factor H repeat homology <FH1>
 F.243-300/Domain: complement factor H repeat homology <FH2>
 F.305-363/Domain: complement factor H repeat homology <FH3>
 F.368-426/Domain: complement factor H repeat homology <FH4>
 F.431-489/Domain: complement factor H repeat homology <FH5>
 F.494-548/Domain: complement factor H repeat homology <FH6>
 F.25-391,528/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 58.1%; Score 126; DB 2; Length 612;
 Best Local Similarity 57.6%; Pred. No. 8,5e-07;
 Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVEIINHTCNCNDVGYGPOCQ 33
 DB 143 CPMSCSGHCEIETINSYCKCHPGLGPNC 175

RESULT 11

JC5092
 E-selectin - pig
 C.Species: Sus scrofa domestica (domestic pig)
 C.Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 19-May-2000
 R.Winkler, H.; Brostjan, C.; Csizmadia, V.; Natarajan, G.; Anrather, J.; Bach, F.H.
 Gene 176, 67-72, 1996
 A.Title: The intron-exon structure of the porcine E-selectin-encoding gene.
 A.Reference number: JC5092; MUID:97075911
 A.Contents: endothelial cells
 A.Accession: JC5092

A.Molecule type: DNA
 A.Residues: 1-482 <MIN>
 A.Cross-references: GB:U07521; NID:91052974; PIDN:AAC48680.1; PID:91052975
 C.Comment: This protein is a member of the selectin family of adhesion molecules.
 C.Genetics:

A.Introns: 13/1; 1421; 178/1; 237/1; 300/1; 363/1; 422/1; 459/1; 466/1
 C.Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement facto
 F.13-139/Domain: C-type lectin homology <CH>
 F.181-233/Domain: complement factor H repeat homology <FH1>
 F.240-298/Domain: complement factor H repeat homology <FH2>
 F.303-361/Domain: complement factor H repeat homology <FH3>
 F.366-420/Domain: complement factor H repeat homology <FH4>

Query Match 58.1%; Score 125; DB 2; Length 482;
 Best Local Similarity 60.6%; Pred. No. 9.1e-07;
 Matches 20; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVEIINHTCNCNDVGYGPOCQ 33
 DB 144 CPMSCSGHCEIETINSYCKCHPGLGPNC 176

RESULT 12

A30359
 P-selectin precursor - human
 N.Alternate names: CD62 antigen; granule membrane protein 140
 C.Species: Homo sapiens (man)
 C.Date: 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 19-May-2000
 C.Accession: A30359
 R.Johnston, G.T.; Cook, R.G.; McEver, R.P.
 Cell 56, 1033-1044, 1989

A.Title: Cloning of GMP-140, a granule membrane protein of platelets and endothelium; se
 A.Reference number: A30359; MUID:89168432
 A.Molecule type: mRNA
 A.Residues: 1-830 <JOH>
 A.Cross-references: GB:M25322
 A.Note: parts of this sequence, including the amino end of the mature protein, were conf
 C.Genetics:
 A.Gene: GDB:SELP; GRMP

A.Cross-references: GDB:120018; OMIM:173610

A.Map position: 1q22-1q25

C.Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;

C.Keywords: cell adhesion; glycoprotein; phospholipid; phospholipid; surface an

F.1-41/Domain: signal sequence #status predicted <SIG>

F.42-830/Product: P-selectin #status experimental <MAT>

F.163-194/Domain: EGF homology <EGF>

F.200-257/Domain: complement factor H repeat homology <FH01>

F.262-319/Domain: complement factor H repeat homology <FH02>

F.324-381/Domain: complement factor H repeat homology <FH03>

F.386-443/Domain: complement factor H repeat homology <FH04>

F.448-505/Domain: complement factor H repeat homology <FH05>

F.510-567/Domain: complement factor H repeat homology <FH06>

F.572-629/Domain: complement factor H repeat homology <FH07>

F.642-699/Domain: complement factor H repeat homology <FH08>

F.704-761/Domain: complement factor H repeat homology <FH09>

F.772-795/Domain: transmembrane #status predicted <TMN>

F.796-830/Domain: intracellular #status predicted <CYT>

F.54,98,180,212,219,411,460,518,665,716,723,741/Binding site: carbohydrate (Asn) (cov

Query Match 58.1%; Score 125; DB 2; Length 830;
 Best Local Similarity 57.6%; Pred. No. 1.4e-06;
 Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVEIINHTCNCNDVGYGPOCQ 33
 DB 163 CQPMSCSGHCEIETINSYCKCHPGLGPNC 195

RESULT 13

P-selectin precursor - mouse
 N.Alternate names: CD62; granule membrane protein 140; PADSEM
 C.Species: Mus musculus (house mouse)
 C.Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 19-May-2000
 C.Accession: A42755; A44899

R.Weller, A.; Isenmann, S.; Vestweber, D.
 J. Biol. Chem. 267, 15176-15183, 1992

A.Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-sele

A.Reference number: A42755; MUID:92340571

A.Accession: A42755

A.Status: nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1-768 <WEL>

A.Cross-references: GB:M87861; NID:9200552; PIDN:AAA40008.1; PID:9200553

A.Experimental source: endothelial cells

A.Note: sequence extracted from NCBI backbone (NCBI:109467)

R.Sanders, W.E.; Wilson, R.W.; Ballantyne, C.M.; Beaudet, A.L.
 Blood 80, 795-800, 1992

A.Title: Molecular cloning and analysis of in vivo expression of murine P-selectin.

A.Reference number: A44899; MUID:92345617

A.Accession: A44899

A.Status: nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1-723, 'E', 725-768 <SAN>

A.Cross-references: GB:M72332; NID:9193565; PIDN:AAA37712.1; PID:9193566

A.Note: sequence extracted from NCBI backbone (NCBI:109900)

C.Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;

C.Keywords: cell adhesion; glycoprotein; phospholipid; phospholipid; transmembr

F.1-11/Domain: signal sequence #status predicted <SIG>

F.42-768/Product: P-selectin #status predicted <MAT>

F.163-194/Domain: EGF homology <EGF>

F.200-257/Domain: complement factor H repeat homology <FH01>

F.262-319/Domain: complement factor H repeat homology <FH02>

F.324-381/Domain: complement factor H repeat homology <FH03>

F.386-443/Domain: complement factor H repeat homology <FH04>

F.448-505/Domain: complement factor H repeat homology <FH05>

F.510-567/Domain: complement factor H repeat homology <FH06>

F.580-637/Domain: complement factor H repeat homology <FH07>

F.642-699/Domain: complement factor H repeat homology <FH08>

F.710-733/Domain: transmembrane #status predicted <TMN>

F.734-768/Domain: intracellular #status predicted <INT>

F:45,54,107,212,347,398,456,467,603,654,661,679/Binding site: carbohydrate (Asn) (covalent)

Query Match 57.7%; Score 124; DB 2; Length 768;
Best Local Similarity 54.5%; Pred. No. 1.7e-06;
Matches 18; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHCEVEIINHTCNCVDVGYGPQCQ 33
Db 163 CQPMSCSGHCEVEIINHTCNCVDVGYGPCE 195

RESULT 14
S36772

E-selectin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 19-May-2000
C:Accession: S36772

R:Nguyen, M.; Strudel, N.A.; Bischoff, J.

Nature 365, 267-269, 1993

A:Title: A role for sialyl Lewis-X/A glycoconjugates in capillary morphogenesis.
A:Reference number: S36772; MUID:93382537

A:Accession: S36772

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-485 <NCU>

A:Cross-references: GB:I12039; NID:g402913; PIDN:AAA02991.1; PID:g402914

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor F:13-139/Domain: C-type lectin homology <LCH>

F:181-237/Domain: complement factor H repeat homology <FH1>

F:242-289/Domain: complement factor H repeat homology <FH2>

F:304-362/Domain: complement factor H repeat homology <FH3>

F:367-421/Domain: complement factor H repeat homology <FH4>

Query Match 54.9%; Score 118; DB 2; Length 485;
Best Local Similarity 54.5%; Pred. No. 5.0e-06;
Matches 18; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHCEVEIINHTCNCVDVGYGPQCQ 33
Db 144 CQPMSCSGHCEVEIINHTCNCVDVGYGPCE 176

RESULT 15
I46709

endothelial leukocyte adhesion molecule 1 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 19-May-2000

C:Accession: I46709; I46708

R:Langran, J.D.; Tsang, T.C.; Rumberger, J.M.; Burns, D.K.

DNA Cell Biol. 11, 149-162, 1992

A:Title: Characterization of cDNA and genomic sequences encoding rabbit ELAM-1: Conservation

A:Reference number: I46708; MUID:92189729

A:Accession: I46709

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-551 <LAR>

A:Cross-references: GB:M91005; NID:g165006; PIDN:AAA1244.1; PID:g165007

A:Accession: I46708

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-307, 'v', 309-327, 'v', 329-490, 'A', 492-551 <LAR2>

A:Cross-references: GB:M91004; NID:g165004; PIDN:AAA1243.1; PID:g165005

C:Genetics:

A:Gene: ELAM1

A:Map position: 1q22-q25

A:Insertions: 13/1; 143/1; 179/1; 241/1; 303/1; 366/1; 429/1; 488/1; 525/1; 533/2

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor

C:Keywords: duplication; glycoprotein; tandem repeat

F:14-140/Domain: C-type lectin homology <LCH>

F:182-239/Domain: complement factor H repeat homology <FH1>

F:244-301/Domain: complement factor H repeat homology <FH2>

F:306-364/Domain: complement factor H repeat homology <FH3>
F:369-427/Domain: complement factor H repeat homology <FH4>
F:432-486/Domain: complement factor H repeat homology <FH5>
F:32,45,201,314,321,466/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 54.0%; Score 116; DB 2; Length 551;
Best Local Similarity 54.5%; Pred. No. 1e-05;
Matches 18; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHCEVEIINHTCNCVDVGYGPQCQ 33
Db 145 CQPMSCSGHCEVEIINHTCNCVDVGYGPCE 177

RESULT 16
I53821

E-selectin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 19-May-2000

C:Accession: I53821

R:Anchamph, J.A.; Oliver, M.G.; Anderson, D.C.; Manning, A.M.

Gene 145, 251-255, 1994

A:Title: Cloning, sequence comparison and in vivo expression of the gene encoding rat

A:Reference number: I53821; MUID:94333817

A:Accession: I53821

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-768 <RES>

A:Cross-references: GB:I23088; NID:g349552; PIDN:AAA0325.1; PID:g349553

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor F:163-158/Domain: C-type lectin homology <LCH>

F:163-194/Domain: EGF homology <EGF>

F:262-319/Domain: complement factor H repeat homology <FHR>

F:510-567/Domain: complement factor H repeat homology <FH06>

F:580-637/Domain: complement factor H repeat homology <FH07>

F:642-699/Domain: complement factor H repeat homology <FH08>

Query Match 50.7%; Score 109; DB 2; Length 768;
Best Local Similarity 48.5%; Pred. No. 8.2e-05;
Matches 16; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHCEVEIINHTCNCVDVGYGPQCQ 33
Db 163 CQPMSCSGHCEVEIINHTCNCVDVGYGPCE 195

RESULT 17
S28941

coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N:Alternate names: Hageman factor

C:Species: Cavia porcellus (guinea pig)

C:Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000

C:Accession: S28941

R:Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe,

Biochim. Biophys. Acta 1159, 113-121, 1992

A:Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage

A:Reference number: S28941; MUID:93003367

A:Accession: S28941

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-603 <SEM>

A:Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA46600.1; PID:g49579

C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology

C:Keywords: hydrolase; serine proteinase

F:46-87/Domain: fibronectin type II repeat homology <IF2>

F:134-169/Domain: fibronectin type I repeat homology <FI1>

F:177-208/Domain: EGF homology <EGF>

F:216-294/Domain: kringle homology <KRG>

F:359-597/Domain: trypsin homology <TRY>


```
Query Match          44.2%; Score 95; DB 2; Length 603;
Best Local Similarity 44.1%; Pred. No. 0.0025;
Matches 15; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPMSCGHGECVEIINHTCMCDVGYGPOCOL 34
DB 177 CQTNPCLNGRGCLFVEGHHLCDCPMGYGPFCDL 210

RESULT 18
T21339
hypotheetical protein F25D7.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21339
R:McMurray, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19409
A:Accession: T21339
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-710 <WIL>
A:Cross-references: EMBL:Z78418; PIDN:CA801698.1; GSPDB:GNO00019; CESP:F25D7.5
A:Experimental source: clone F25D7
C:Genetics:
A:Gene: CESP:F25D7.5
A:Map position: 1
A:introns: 20/1; 56/2; 87/1; 137/3; 235/3; 274/2; 410/2; 473/2; 496/3

Query Match          44.2%; Score 95; DB 2; Length 710;
Best Local Similarity 38.2%; Pred. No. 0.0029;
Matches 13; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 1 COPMSCGHGECVEIINHTCMCDVGYGPOCOL 34
DB 148 CPTTCNGHCKCYDVEDVKDCYWGVEGHECEV 181

RESULT 19
Xotch protein - African clawed frog
A35844
C:Species: Xenopus laevis (African clawed frog)
C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 13-Aug-1999
C:Accession: A35844
R:Cotfman, C.; Harris, W.; Kintner, C.
Science 249, 1438-1441, 1990
A:Title: Xotch, the Xenopus homolog of Drosophila notch.
A:Reference number: A35844; MUID:90385285
A:Accession: A35844
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-2524 <COF>
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C:Keywords: transmembrane protein
F:146-177/Domain: EGF homology <EGX1>
F:184-215/Domain: EGF homology <EGF1>
F:222-254/Domain: EGF homology <EGF>
F:456-487/Domain: EGF homology <EGX2>
F:757-788/Domain: EGF homology <EGF3>
F:1025-1056/Domain: EGF homology <EGX3>
F:1924-1956/Domain: ankyrin repeat homology <AN1>
F:1957-1989/Domain: ankyrin repeat homology <AN2>
F:1991-2023/Domain: ankyrin repeat homology <AN3>
F:2024-2056/Domain: ankyrin repeat homology <AN4>
F:2057-2089/Domain: ankyrin repeat homology <AN5>

Query Match          43.3%; Score 93; DB 2; Length 2524;
Best Local Similarity 45.5%; Pred. No. 0.013;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPMSCGHGECVEIINHTCMCDVGYGPOCOL 33

DB 911 COPNCHNGSCSDGINMFNCMPGRFRPKCE 943

RESULT 20
T31070
notch homolog - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C:Accession: T31070
R:Sherwood, D.R.; McClay, D.R.
Development 124, 3363-3374, 1997
A:Title: Identification and localization of a sea urchin Notch homologue: Insights in
A:Reference number: Z20966; MUID:97454256
A:Accession: T31070
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2531 <SHE>
A:Cross-references: EMBL:AF000634; NID:92570350; PID:92570351; PIDN:AA882088.1
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match          41.9%; Score 90; DB 2; Length 2531;
Best Local Similarity 38.2%; Pred. No. 0.029;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPMSCGHGECVEIINHTCMCDVGYGPOCOL 34
DB 555 COSPCENGTCIDGVOFTCICETGVEGRCEM 588

RESULT 21
S42373
hypotheetical protein T2065.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S42373
R:Smith, A.
submitted to the EMBL Data Library, March 1994
A:Reference number: S42368
A:Accession: S42373
A:Molecule type: DNA
A:Residues: 1-3051 <SMT>
A:Cross-references: EMBL:Z30423; NID:9458479; PID:9458485
C:Genetics:
A:introns: 96/1; 166/3; 201/2; 355/1; 404/1; 427/2; 794/1; 892/1; 952/2; 1041/1; 1146
C:Superfamily: von Willebrand factor type A repeat homology; EGF homology; fibronecti
F:512-679/Domain: von Willebrand factor type A repeat homology <WMA1>
F:754-793/Domain: fibronectin type II repeat homology <2FI>
F:1201-1244/Domain: EGF homology <EGF>

Query Match          41.9%; Score 90; DB 2; Length 3051;
Best Local Similarity 51.6%; Pred. No. 0.034;
Matches 16; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

OY 6 CSGHGEVVE--IINHTCMCDVGYGPOCOL 34
DB 2664 CNAHGDVCHNFTATNNTICVCTDGTGWTGPOCOV 2694

RESULT 22
A49175
Notch B protein - mouse (fragment)
N:Alternate names: Notch homolog
C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-1994 #sequence_revision 05-Jan-1996 #text_change 20-Sep-1999
C:Accession: A49175; PH1570; S32113
R:Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A:Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety
A:Reference number: A49175; MUID:93178563
A:Accession: A49175
```

A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1203 <L2AR>
A:Cross-references: EMBL:X68279; NID:g287989; PIDN:CAA8340.1; PID:g287990
A:Experimental source: embryo
A>Note: sequence extracted from NCBI backbone (NCBI:126158)
C:Comment: This protein has many EGF repeats and 11n-12/Notch repeats.
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:143-174/Domain: EGF homology <EGX1>
F:482-513/Domain: EGF homology <EGF1>
F:560-591/Domain: EGF homology <EGF2>
F:674-705/Domain: EGF homology <EGX2>
F:712-743/Domain: EGF homology <EGF3>
F:836-867/Domain: EGF homology <EGX3>

Query Match 41.4%; Score 89; DB 2; Length 1203;
Best Local Similarity 41.2%; Pred. No. 0.021;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHCEVETIINHTNCNDVGYGPOCQ 34
Db 181 CQSNPCVNNGCVDKVNRFQCLCPGFTGPVCQI 214

RESULT 23
notch3 protein - human

C:Species: Homo sapiens (man)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 20-Sep-1999
C:Accession: S78549; S71825
R:Joutel, A.; Tournier-Lasserre, E.
submitted to the EMBL Data Library, April 1997
A:Reference number: S78549
A:Accession: S78549
A:Molecule type: mRNA
A:Residues: 1-2321 <JOU1>
A:Cross-references: EMBL:097669; NID:g2668591; PIDN:AAB91371.1; PID:g2668592
R:Joutel, A.; Corpechot, C.; Ducros, A.; Vahedi, K.; Chabrier, H.; Mouton, P.; Alamowicz
X, M.M.; Weissbach, J.; Bach, J.F.; Bousser, M.G.; Tournier-Lasserre, E.
Nature 383, 707-710, 1996
A:Title: Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke
A:Reference number: S71825; MUID:97032728
A:Accession: S71825
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 67-113;138-194;268-333; 'G',335-346;536-613;716-765;1240-1279;1815-1888 <JOUT2
A:Cross-references: EMBL:097669
C:Genetics:
A:Gene: notch3
A:Map position: 19p13.1
C:Function:
A:Description: may be involved in pathogenesis of CADASIL, causing a type of stroke and
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C:Keywords: tandem repeat; transmembrane protein
F:123-155/Domain: EGF homology <EGX1>
F:162-194/Domain: EGF homology <EGF1>
F:240-271/Domain: EGF homology <EGX2>
F:318-349/Domain: EGF homology <EGF>
F:473-504/Domain: EGF homology <EGX3>
F:553-584/Domain: EGF homology <EGF3>
F:928-959/Domain: EGF homology <EGX4>
F:1070-1126/Domain: laminin-type EGF-like homology <LEG>
F:1838-1870/Domain: ankyrin repeat homology <AN1>
F:1871-1903/Domain: ankyrin repeat homology <AN2>
F:1905-1937/Domain: ankyrin repeat homology <AN3>
F:1938-1970/Domain: ankyrin repeat homology <AN4>
F:1971-2003/Domain: ankyrin repeat homology <AN5>

Query Match 41.4%; Score 89; DB 2; Length 2321;
Best Local Similarity 45.5%; Pred. No. 0.035;

Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHCEVETIINHTNCNDVGYGPOCQ 33
Db 928 CSPSCFNGGTCDVGVNSFSCLCRPGYGAHCQ 960

RESULT 24

A49128
cell fate determining gene Notch2 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C:Accession: A49128
R:Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 116, 931-941, 1992
A:Title: Notch2: a second mammalian Notch gene.
A:Reference number: A49128; MUID:93202015
A:Accession: A49128
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2471 <ME1>
A:Experimental source: Schwann cell
A>Note: sequence extracted from NCBI backbone (NCBI:127811)
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:264-295/Domain: EGF homology <EGX1>
F:799-830/Domain: EGF homology <EGF1>
F:877-908/Domain: EGF homology <EGX2>
F:1029-1060/Domain: EGF homology <EGF>
F:1067-1098/Domain: EGF homology <EGX3>
F:1153-1184/Domain: EGF homology <EGF3>
F:1191-1222/Domain: EGF homology <EGX4>
F:1876-1908/Domain: ankyrin repeat homology <AN1>
F:1909-1941/Domain: ankyrin repeat homology <AN2>
F:1943-1975/Domain: ankyrin repeat homology <AN3>
F:1976-2008/Domain: ankyrin repeat homology <AN4>
F:2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 41.4%; Score 89; DB 2; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.037;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHCEVETIINHTNCNDVGYGPOCQ 34
Db 498 CQSNPCVNNGCVDKVNRFQCLCPGFTGPVCQI 531

RESULT 25

JX0210
protein C (activated) (EC 3.4.21.69) precursor - mouse
N:Alternate names: vitamin K-dependent serine protease
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JX0210
R:Rade, N.; Sato, M.; Tsujimura, A.; Iwase, R.; Hashimoto-Gotoh, T.
J. Biochem. 111, 491-495, 1992
A:Title: Isolation and characterization of a mouse protein C cDNA.
A:Reference number: JX0210; MUID:92316897
A:Accession: JX0210
A:Molecule type: mRNA
A:Residues: 1-461 <TAD>
A:Cross-references: GB:D10445; NID:g220385; PIDN:BAA01235.1; PID:g220386
A:Experimental source: liver
C:Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that
s.
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
F:1-33/Domain: signal sequence #status predicted <SIG>
F:27-85/Domain: Gla domain homology <Gla>
F:34-41/Domain: propeptide #status predicted <PRO>
F:42-196,199-461/Product: protein C #status predicted <PRC>
F:42-196/Domain: light chain #status predicted <PCL>
F:91-130/Domain: EGF homology <EG1>

```
F:139-174/Domain: EGF homology <EG2>
F:199-461/Domain: heavy chain #status predicted <PCH>
F:199-211/Domain: activation peptide #status predicted <ACT>
F:212-461/Product: vitamin K dependent serine proteolase #status predicted <VIT>
F:212-445/Domain: trypsin homology <TRY>
F:47-48,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F:121-130,139-150,146-159,161-174,182-319,238-254,373-387,398-426/Disulfide bonds: #stat
F:214,290,355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:253,299,402/Active site: His, Asp, Ser #status predicted

Query Match          40.5%; Score 87; DB 1; Length 461;
Best Local Similarity 46.4%; Pred. No. 0.016;
Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 6 CSHGECVEITNNHNCNDVGYGPOCQ 33
DB 104 CCGHGTICIDIGISFCSCDKGMEGKFCQ 131

RESULT 26
150719
C-Delta-1 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I50719
R:Hemilique, D.; Adam, J.; Myat, A.; Chitnis, A.; Lewis, J.; Ish-Horowicz, D.
Nature 375, 787-790, 1995
A:Title: Expression of a Delta homologue in prospective neurons in the chick.
A:Reference number: I50719; MUID:95319507
A:Accession: I50719
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-728 <HEN>
A:Cross-references: EMBL:026590; NID:9882411; PIDN:AAC59689.1; PID:9882412
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:299-333/Domain: EGF homology <EGX1>
F:339-370/Domain: EGF homology <EGF1>
F:416-447/Domain: EGF homology <EGX2>
F:454-485/Domain: EGF homology <EGF>
F:492-523/Domain: EGF homology <EGF3>

Query Match          40.5%; Score 87; DB 2; Length 728;
Best Local Similarity 35.3%; Pred. No. 0.023;
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 1 COPWSCGHGECVEITNNHNCNDVGYGPOCQ 34
DB 339 CDANPCKNGSCTDLENSTYCTCPGFGKNCLEL 372

RESULT 27
A40136
fibropellin Ia - sea urchin (Strongylocentrotus purpuratus)
N:Alternate names: epidermal growth factor homolog precursor
M:Contains: alternatively spliced fibropellin Ib (EGF1)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 13-May-1992 #sequence_revision 17-Sep-1997 #text_change 21-Jul-2000
C:Accession: A40136; A40136; A29316; A29316; A43131
R:Delgadillo-Reynoso, M.G.; Rollo, D.R.; Hursh, D.A.; Raff, R.A.
J Mol. Evol. 29, 314-327, 1989
A:Title: Structural analysis of the uEGF gene in the sea urchin Strongylocentrotus purp
A:Reference number: A40136; MUID:90112459
A:Accession: A40136
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-114 <DEL>
A:Cross-references: GB:X17530; NID:910225; PID:9667061
A:Accession: B40136
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA

A:Residues: 181-251,329-370,'R',372-408,'RA',411-441 <DE2>
A:Accession: C40136
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 'K',747-821,898-978 <DE3>
R:Hursh, D.A.; Andrews, M.E.; Raff, R.A.
Science 237, 1487-1490, 1987
A:Title: A sea urchin gene encodes a polypeptide homologous to epidermal growth facto
A:Reference number: A29316; MUID:87319677
A:Accession: A29316
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 'S',280-481,786-1064 <HUR>
A:Cross-references: GB:M17421; NID:9161474; PIDN:AAA30050.1; PID:9552260
R:Hunt, L.T.; Barker, W.C.
FASEB J. 3, 1760-1764, 1989
A:Title: Avidin-like domain in an epidermal growth factor homolog from a sea urchin.
A:Reference number: A43131; MUID:89196806
A:Contents: annotation
C:Comment: EGF homology repeats 10-17 are spliced out in the short form (fibropellin
C:Superfamily: C1r/C1s repeat homology; EGF homology
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1064/Product: fibropellin I #status predicted <PIB>
F:23-54/Domain: EGF homology <EG01>
F:57-175/Domain: C1r/C1s repeat homology <C1R>
F:180-211/Domain: EGF homology <EG02>
F:218-249/Domain: EGF homology <EG03>
F:256-287/Domain: EGF homology <EG04>
F:294-325/Domain: EGF homology <EG05>
F:332-363/Domain: EGF homology <EG06>
F:370-401/Domain: EGF homology <EG07>
F:408-439/Domain: EGF homology <EG08>
F:446-477/Domain: EGF homology <EG09>
F:484-515/Domain: EGF homology <EG10>
F:522-553/Domain: EGF homology <EG11>
F:560-591/Domain: EGF homology <EG12>
F:598-629/Domain: EGF homology <EG13>
F:636-667/Domain: EGF homology <EG14>
F:674-705/Domain: EGF homology <EG15>
F:712-743/Domain: EGF homology <EG16>
F:750-781/Domain: EGF homology <EG17>
F:788-819/Domain: EGF homology <EG18>
F:826-857/Domain: EGF homology <EG19>
F:864-895/Domain: EGF homology <EG20>
F:902-933/Domain: EGF homology <EG21>
F:936-1064/Region: avidin-like
F:23-34,28-43,45-54,62-88,180-191,185-200,202-211,218-229,223-238,240-249,256-267,261
57,451-466,468-477,484-495/Disulfide bonds: #status predicted
F:489-504,506-515,522-533,527-543,544-553,560-571,565-580,582-591,598-609,603-618,620
06,810-819,826-837,831-846,848-857,864-875,869-884,886-895,902-913,907-922,924-933/DI

Query Match          40.5%; Score 87; DB 2; Length 1064;
Best Local Similarity 39.4%; Pred. No. 0.031;
Matches 13; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 COPWSCGHGECVEITNNHNCNDVGYGPOCQ 33
DB 750 CASMPLNGACAGIEMVNGYTCOCVAGYGVICE 782

RESULT 28
A35672
crumbs protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 11-Jan-2000
C:Accession: A35672
R:Teppas, U.; Theres, C.; Knust, E.
Cell 61, 787-799, 1990
A:Title: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophi
A:Reference number: A35672; MUID:90263104
A:Accession: A35672
A:Status: preliminary
```


A:Cross-references: EMBL:U24703; NID:g902486; PID:g902487
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 F:1-27/Domain: signal sequence #status predicted <Sig>
 F:28-3461/Product: reelin #status predicted <MAT>
 F:1769-1795/Domain: EGF homology <EGF>

Query Match 39.8%; Score 85.5; DB 2; Length 3461;
 Best Local Similarity 51.7%; Pred. No. 0.12;
 Matches 15; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 6 CSGHGECEIINNNHTCNCVGYGPOCQ 34
 DB 2137 CYGHGSC---INGTKICDPGYSGPTCKI 2162

RESULT 32
 E89753
 Protein F11C7.4 [Imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
 C:Accession: E89753
 R:Anonymous, The C. elegans Sequencing Consortium.
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: E89753
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1722 <STO>
 A:Cross-references: GB:chr_X; PIDN:AAC69012.1; PID:g1125776; GSPDB:GN00028; CESP:F11C7.4
 C:Genetics:
 A:Gene: F11C7.4
 A:Map position: X

Query Match 39.5%; Score 85; DB 2; Length 1722;
 Best Local Similarity 42.4%; Pred. No. 0.077;
 Matches 14; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNNHTCNCVGYGPOCQ 33
 DB 1371 CSSNTCSSRGACSPVWNNTVYCNCDNMWRAHQ 1403

RESULT 33
 T30201
 Notch homolog protein - sea squirt (Halocynthia roretzi)
 C:Species: Halocynthia roretzi
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C:Accession: T30201
 R:Hoti, S.; Saitoh, T.; Matsumoto, M.; Makabe, K.W.; Nishida, H.
 Dev. Genes Evol. 207, 371-380, 1997
 A:Title: Notch homologue from Halocynthia roretzi is preferentially expressed in the cer
 A:Reference number: 220775
 A:Accession: T30201
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2352 <HOR>
 A:Cross-references: EMBL:AB001327; NID:d1204472; PID:d1026501; PIDN:BA25571.1
 C:Genetics:
 A:Gene: Notch

Query Match 39.5%; Score 85; DB 2; Length 2352;
 Best Local Similarity 43.8%; Pred. No. 0.1;
 Matches 14; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNNHTCNCVGYGPOC 32
 DB 113 CSPNCSNGACEELNSFKCTGSGYGDTC 144

RESULT 34
 T29359
 Hypothetical protein R05G6.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29359
 R:Murray, J.; Le, T.T.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid R05G6.
 A:Reference number: Z20612
 A:Accession: T29359

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-372 <MUR>
 A:Cross-references: EMBL:U58746; PIDN:AAB00626.1; GSPDB:GN00022; CESP:R05G6.9
 A:Experimental source: strain Bristol N2; clone R05G6
 C:Genetics:
 A:Gene: CESP:R05G6.9
 A:Map position: 4
 A:introns: 80/1; 161/1; 245/1; 286/1

Query Match 39.1%; Score 84; DB 2; Length 372;
 Best Local Similarity 40.7%; Pred. No. 0.029;
 Matches 11; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 6 CSGHGECEIINNNHTCNCVGYGPOC 32
 DB 339 CSGNGICISLKSYSCTCNLGMTGPTC 365

RESULT 35
 J01322
 tenascin precursor - mouse
 N:Alternate names: contractin; hexabrachion
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: J01322; A37936; B37936; S14571; S50209
 R:Saga, Y.; Tsukamoto, T.; Jing, N.; Kusakabe, M.; Sakakura, T.
 Gene 104, 177-185, 1991
 A:Title: Murine tenascin: cDNA cloning, structure and temporal expression of isoforms
 A:Reference number: J01322; MUID:92009211
 A:Accession: J01322
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-2019 <SAG>
 A:Cross-references: GB:D90343; NID:9220609; PIDN:BA14355.1; PID:9220610
 A:Experimental source: cell line 2H6GR
 A:Note: the authors translated the codon ATG for residue 60 as Trp
 R:Weller, A.; Beck, S.; Ekblom, P.
 J. Cell Biol. 112, 355-362, 1991
 A:Title: Amino acid sequence of mouse tenascin and differential expression of two ten
 A:Reference number: A37936; MUID:91107734
 A:Accession: A37936
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 10
 A:Cross-references: GB:X56304
 A:Accession: B37936
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 10
 A:Cross-references: GB:X56304
 R:Weller, A.; Beck, S.; Ekblom, P.
 submitted to the EMBL Data Library, August 1990
 A:Description: Amino acid sequence of mouse tenascin and differential expression of tw
 A:Reference number: S14571
 A:Accession: S14571
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-201, 'E', 203-317, 'S', 319-1018, 'S', 1020-1024, 'H', 1026-1305, 'S', 1307-2019

A:Cross-references: EMBL:X56304; NID:g54768; PIDN:CAA39751.1; PID:g54769
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by
A:Reference number: S50206; MUID:95035091
A:Accession: S50209
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 46-146 <GLD>
A:Cross-references: EMBL:X80281
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type
C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracellular
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-2019/Product: tenascin, long splice form #status predicted <MAT>
F:23-2019/Product: tenascin, short splice form #status predicted <MAT2>
F:23-1071,1527-2019/Product: tenascin, short splice form #status predicted <MAT2>
F:408-434/Domain: EGF homology <EGF>
F:622-703/Domain: fibronectin type III repeat homology <FN3A>
F:711-793/Domain: fibronectin type III repeat homology <FN3B>
F:802-884/Domain: fibronectin type III repeat homology <FN3C>
F:892-976/Domain: fibronectin type III repeat homology <FN3D>
F:984-1064/Domain: fibronectin type III repeat homology <FN3E>
F:1073-1155/Domain: fibronectin type III repeat homology <FN3F>
F:1346-1428/Domain: fibronectin type III repeat homology <FN3G>
F:1437-1519/Domain: fibronectin type III repeat homology <FN3H>
F:1527-1608/Domain: fibronectin type III repeat homology <FN3I>
F:1616-1696/Domain: fibronectin type III repeat homology <FN3J>
F:1704-1784/Domain: fibronectin type III repeat homology <FN3K>
F:1799-2007/Domain: fibrinogen beta/gamma homology <FBG>
F:38,166,184,327,788,1018,1079,1093,1119,1184,1210,1275,1301,1354,1364,1394,1627,1878,19

Query Match 39.1%; Score 84; DB 1; Length 2019;
Best Local Similarity 53.3%; Pred. No. 0.11;
Matches 16; Conservative 0; Mismatches 10; Indels 4; Gaps 1;

Qy 3 PWSGSHGECVEIINHTCNCNVGYGPGC 32
Db 409 PNCSCGHGRCV----NGCVCDEGTGEDC 434

RESULT 36
A32160
tenascin-C - human
N:Alternate names: hexabrachion
C:Species: Homo sapiens (man)
C:Date: 31-Jul-1989 #sequence_revision 12-Apr-1996 #text_change 20-Aug-1999
C:Accession: I38337; A32160; S14015; S16166; S50208; S49354
R:Gherzi, R.; Carremolla, B.; Sirtl, A.; Fonassi, M.; Balza, E.; Zardi, L.
J. Biol. Chem. 270, 3429-3434, 1995
A:Title: Human tenascin gene: structure of the 5'-region, identification, and character
A:Accession: I38337
A:Reference number: A55974; MUID:95155442
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2201 <RES>
A:Cross-references: EMBL:X78565; NID:g556844; PIDN:CAA55309.1; PID:g556845
R:Gulcher, J.R.; Nies, D.E.; Marton, L.S.; Stefansson, K.
Proc. Natl. Acad. Sci. U.S.A. 86, 1588-1592, 1989
A:Title: An alternatively spliced region of the human hexabrachion contains a repeat of
A:Reference number: A32160; MUID:85160821
A:Accession: A32160
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 431-538 'R', 540-1065 'H', 1067-1599, 'LMLHPRASN', 1609-2054, 'L' <GUL>
A:Cross-references: GB:M24630; NID:g514363; PIDN:AAA52703.1; PID:g553348
R:Sirtl, A.; Carremolla, B.; Segineli, M.; Leprini, A.; Casari, G.; Baralle, F.; Zardi, L.
Nucleic Acids Res. 19, 525-531, 1991
A:Title: Human tenascin: primary structure, pre-mRNA splicing patterns and localization
A:Reference number: S14015; MUID:91187670
A:Accession: S14015
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-243,245-369, 'V', 371-679, 'R', 681-1676, 'I', 1678-2139, 'TR', 2142,2144-2201
A:Cross-references: EMBL:X56160; NID:g37226; PIDN:CAA39628.1; PID:g37227
R:Nies, D.E.; Hemeseth, T.J.; Kim, J.H.; Gulcher, J.R.; Stefansson, K.
J. Biol. Chem. 266, 2818-2823, 1991
A:Title: The complete cDNA sequence of human hexabrachion (tenascin). A multidomain p
A:Reference number: S16166; MUID:91131572
A:Accession: S16166
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-538 'R', 540-1065 'H', 1067-1599, 'LMLHPRASN', 1609-2054, 'LH', 2055-2201 <NI
A:Cross-references: EMBL:M55618; NID:g184483; PIDN:AAA88083.1; PID:g184484
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues
A:Reference number: S50206; MUID:95035091
A:Accession: S50208
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 46-125 <GLD>
A:Cross-references: EMBL:X80280
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Gene: GDB:HXB
A:Cross-references: GDB:120073; OMIM:187380
A:Map position: 9q33-q9q33
C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin ty
C:Keywords: alternative splicing; extracellular matrix
F:622-703/Domain: fibronectin type III repeat homology <FN3A>
F:711-794/Domain: fibronectin type III repeat homology <FN3B>
F:892-976/Domain: fibronectin type III repeat homology <FN3C>
F:1437-1519/Domain: fibronectin type III repeat homology <FN3D>
F:1619-1701/Domain: fibronectin type III repeat homology <FN3E>
F:1709-1790/Domain: fibronectin type III repeat homology <FN3F>
F:1798-1878/Domain: fibronectin type III repeat homology <FN3G>
F:1886-1966/Domain: fibronectin type III repeat homology <FN3H>
F:1981-2189/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 39.1%; Score 84; DB 2; Length 2201;
Best Local Similarity 53.3%; Pred. No. 0.12;
Matches 16; Conservative 0; Mismatches 10; Indels 4; Gaps 1;

Qy 3 PWSGSHGECVEIINHTCNCNVGYGPGC 32
Db 409 PNCSCGHGRCV----NGCVCDEGTGEDC 434

RESULT 37
S45306
notch 3 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
C:Accession: S45306
R:Riandelli, M.; Dahlstrand, J.; Lendahl, U.
Mech. Dev. 46, 123-136, 1994
A:Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth fact
A:Reference number: S45306; MUID:95001556
A:Accession: S45306
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2318 <LAR>
A:Cross-references: EMBL:X74760; NID:g483580; PIDN:CAA52276.1; PID:g483581
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:163-195/Domain: EGF homology <EGF1>
F:474-505/Domain: EGF homology <EGF2>
F:854-885/Domain: EGF homology <EGF3>
F:1839-1871/Domain: ankyrin repeat homology <AN1>
F:1872-1904/Domain: ankyrin repeat homology <AN2>
F:1906-1938/Domain: ankyrin repeat homology <AN3>
F:1939-1971/Domain: ankyrin repeat homology <AN4>
F:1972-2004/Domain: ankyrin repeat homology <AN5>

Query Match 39.1%; Score 84; DB 2; Length 2318;
Best Local Similarity 42.4%; Pred. No. 0.13;
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHGECVEIINNHNTCMCDVGYGPOC 33
DB 929 CSPSSCFNGTCVDCVSSFSCLCRPGYTGTTCO 961

RESULT 38

T09456
Intrinsic factor-B12 receptor Cubillin precursor - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09456
R:Koziyal, R.; Kristiansen, M.; Sillahtaroglu, A.; Hansen, C.; Jacobsen, C.; Tommerup, N.
Blood 91, 3593-3600, 1998
A:Title: The human intrinsic factor-vitamin B12 receptor, cubillin: Molecular characterization.
A:Reference number: Z16677; MUID:98241400
A:Accession: T09456
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3623 <KOZ>
A:Cross-references: EMBL:AF034611; NID:g3929528; PIDN:AAC82612.1; PID:g3929529
C:Genetics:
A:Map position: 10p12
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: receptor; vitamin B12 uptake
F:1-24/Domains: signal sequence #status predicted <SIG>
F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>
F:436-467/Domains: EGF homology <EGF>

Query Match 39.1%; Score 84; DB 2; Length 3623;
Best Local Similarity 43.3%; Pred. No. 0.18;
Matches 13; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

OY 3 PMSCSGHGECVEIINNHNTCMCDVGYGPOC 32
DB 181 PLSCONGTCVNTMGSYSCCHPEYGPQC 210

RESULT 39

T42681
hypothetical protein DKFzP434E0321.1 - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T42681
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z22233
A:Accession: T42681
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-870,871-1069 <AAA>
A:Cross-references: EMBL:AL133021
A:Experimental source: adult testis; clone DKFzP434E0321
A:Note: the cDNA sequence contains a +1 frameshift near codon 870
C:Genetics:
A:Note: DKFzP434E0321.1

Query Match 38.8%; Score 83.5; DB 2; Length 1069;
Best Local Similarity 42.4%; Pred. No. 0.078;
Matches 14; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

OY 1 COPWSCGHGECVEIINNHNTCMCDVGYGPOC 32
DB 563 CLPGCSDBHGQCDGRTGSGQCLCTGTGTGTPSC 595

RESULT 40

T13953
MEGF5 protein - rat
N:Alternate names: slit protein homolog
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13953
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs.
A:Reference number: Z14126; MUID:98360089
A:Accession: T13953
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1523 <NAK>
A:Cross-references: EMBL:AB011531; NID:g3449291; PIDN:BA32461.1; PID:g3449292
C:Genetics:
A:Gene: MEGF5

Query Match 38.8%; Score 83.5; DB 2; Length 1523;
Best Local Similarity 50.0%; Pred. No. 0.1;
Matches 16; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

OY 1 COPWSCGHGECVEIINNHNTCMCDVGYGPOC 32
DB 1372 CLGHSCS-HGTCVANGNSYVCKACBEGEPLC 1402

Search completed: September 7, 2002, 10:15:28
Job time: 206 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:22:08 : Search time 85.16 Seconds
(without alignments)
100.974 Million cell updates/sec

Title: US-09-119-209-2_COPY_160_193
Perfect score: 215
Sequence: 1 CQPMSCSGHGECEIINNHTCNCVDVGYGPOCOL 34

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 869228 seqs, 252910224 residues

Total number of hits satisfying chosen parameters: 869228

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	100.0	372	5	US-09-791-537-152667
2	209	97.2	372	5	US-09-791-537-84593
3	205	95.3	323	5	US-09-791-537-53485
4	205	95.3	341	6	US-10-211-364-1171
5	205	95.3	341	6	US-10-212-054-1528
6	205	95.3	341	6	US-10-212-778-1158
7	205	95.3	363	5	US-09-791-537-58446
8	205	95.3	372	5	US-09-791-537-22816
9	205	95.3	372	5	US-09-791-537-42655
10	205	95.3	372	5	US-09-791-537-42657
11	205	95.3	372	5	US-09-791-537-42658
12	205	95.3	372	5	US-09-791-537-42659
13	205	95.3	372	5	US-09-791-537-65658
14	205	95.3	385	5	US-09-791-537-51391
15	205	95.3	385	5	US-09-791-537-53844
16	205	95.3	385	5	US-09-791-537-113060
17	191	88.8	370	5	US-09-791-537-50403
18	181	84.2	376	5	US-09-791-537-132144
19	169	78.6	360	5	US-09-791-537-65053
20	169	78.6	372	5	US-09-791-537-37750
21	159	74.0	372	5	US-09-791-537-6693
22	159	74.0	372	5	US-09-791-537-81233
23	131	60.9	611	5	US-09-791-537-118878
24	130	60.5	162	5	US-09-791-537-23269
25	130	60.5	610	1	PCT-US02-23913-357
26	130	60.5	610	5	US-09-791-537-22813

27	130	60.5	610	5	US-09-791-537-121834	Sequence 121834,
28	130	60.5	610	6	US-10-205-823-357	Sequence 357, App
29	127	59.1	549	5	US-09-791-537-50406	Sequence 50406, A
30	127	59.1	646	5	US-09-791-537-84829	Sequence 84829, A
31	127	59.1	649	5	US-09-791-537-93873	Sequence 93873, A
32	127	59.1	769	5	US-09-791-537-50409	Sequence 50409, A
33	126	58.6	612	5	US-09-791-537-73416	Sequence 73416, A
34	126	58.6	618	5	US-09-791-537-81892	Sequence 81892, A
35	126	58.6	619	5	US-09-791-537-33749	Sequence 33749, A
36	125	58.1	40	5	US-09-791-537-86308	Sequence 86308, A
37	125	58.1	482	5	US-09-791-537-41672	Sequence 41672, A
38	125	58.1	482	5	US-09-791-537-111285	Sequence 111285, A
39	125	58.1	484	5	US-09-791-537-50405	Sequence 50405, A
40	125	58.1	616	5	US-09-791-537-332025	Sequence 332025, A
41	125	58.1	740	5	US-09-791-537-332023	Sequence 332023, A
42	125	58.1	740	5	US-09-791-537-32024	Sequence 32024, A
43	125	58.1	830	5	US-09-791-537-22819	Sequence 22819, A
44	125	58.1	830	5	US-09-791-537-35618	Sequence 35618, A
45	124	57.7	768	5	US-09-791-537-20989	Sequence 20989, A

ALIGNMENTS

```
RESULT 1
US-09-791-537-152667
; Sequence 152667, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 152667
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-152667

Query Match      100.0%  Score 215;  DB 5;  Length 372;
Best Local Similarity 100.0%;  Pred. No. 4e-16;
Matches 34;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      1  CQPMSCSGHGECEIINNHTCNCVDVGYGPOCOL 34
        ||||||||||||||||||||||||||||
DB      160 CQPMSCSGHGECEIINNHTCNCVDVGYGPOCOL 193

RESULT 2
US-09-791-537-84593
; Sequence 84593, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84593
```

```
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-84593
```

```
Query Match          97.2%; Score 209; DB 5; Length 372;
Best Local Similarity 97.1%; Pred. No. 1,8e-15;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 34
DB 160 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 193
```

```
RESULT 3
US-09-791-537-53485
Sequence 53485, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 53485
LENGTH: 323
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-53485
```

```
Query Match          95.3%; Score 205; DB 5; Length 323;
Best Local Similarity 97.0%; Pred. No. 4.5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33
DB 173 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 205
```

```
RESULT 4
US-10-211-364-1171
Sequence 1171, Application US/10211364
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P216CIN
CURRENT APPLICATION NUMBER: US/10/211,364
CURRENT FILING DATE: 2002-08-05
PRIOR APPLICATION NUMBER: 09/760,486
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
```

```
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1778
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1171
LENGTH: 341
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (215)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-211-364-1171
```

```
Query Match          95.3%; Score 205; DB 6; Length 341;
Best Local Similarity 97.0%; Pred. No. 4.7e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33
DB 182 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 214
```

```
RESULT 5
US-10-212-054-1328
Sequence 1328, Application US/10212054
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P212CIN
CURRENT APPLICATION NUMBER: US/10/212,054
CURRENT FILING DATE: 2002-08-06
NUMBER OF SEQ ID NOS: 2164
Prior application removed - See File Wrapper or PALM
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1328
LENGTH: 341
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (215)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-054-1328
```

```
Query Match          95.3%; Score 205; DB 6; Length 341;
Best Local Similarity 97.0%; Pred. No. 4.7e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33
DB 182 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 214
```

```
RESULT 6
US-10-212-778-1158
Sequence 1158, Application US/10212778
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PM026CIN
CURRENT APPLICATION NUMBER: US/10/212,778
CURRENT FILING DATE: 2002-08-07
PRIOR APPLICATION NUMBER: 09/758,449
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
NUMBER OF SEQ ID NOS: 1478
SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 1158
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-778-1158
```

```
Query Match          95.3%; Score 205; DB 6; Length 341;
Best Local Similarity 97.0%; Pred. No. 4.7e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 182 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 214
```

```
RESULT 7
US-09-791-537-58446
; Sequence 58446, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58446
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-58446
```

```
Query Match          95.3%; Score 205; DB 5; Length 363;
Best Local Similarity 97.0%; Pred. No. 4.9e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 173 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 205
```

```
RESULT 8
US-09-791-537-22816
; Sequence 22816, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22816
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-22816
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
```

```
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
RESULT 9
US-09-791-537-42655
; Sequence 42655, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42655
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-791-537-42655
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
RESULT 10
US-09-791-537-42657
; Sequence 42657, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42657
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pan troglodytes
US-09-791-537-42657
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
RESULT 11
US-09-791-537-42658
; Sequence 42658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
```

```

; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Papio hamadryas
US-09-791-537-42658
```

```

Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 33
Db 160 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 192
```

```

RESULT 12
; Sequence 42659, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42659
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pongo pygmaeus
US-09-791-537-42659
```

```

Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 33
Db 160 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 192
```

```

RESULT 13
; Sequence 69658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-69658
```

```

Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 33
Db 160 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 192
```

```

RESULT 14
; Sequence 51391, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51391
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-51391
```

```

Query Match          95.3%; Score 205; DB 5; Length 385;
Best Local Similarity 97.0%; Pred. No. 5.2e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 33
Db 173 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 205
```

```

RESULT 15
; Sequence 53844, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53844
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53844
```

```

Query Match          95.3%; Score 205; DB 5; Length 385;
Best Local Similarity 97.0%; Pred. No. 5.2e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 33
Db 173 CQPMSCGHGECVEIINNHTCNCVDVGYGPQCQ 205
```

```

RESULT 16
US-09-791-537-113060
```

```
; Sequence 113060, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113060
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-113060

Query Match          95.3%; Score 205; DB 5; Length 385;
Best Local Similarity 97.0%; Pred. No. 5.2e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 COPMSCSGHCEVEIINNHTCNCDDVGYGPOCQ 33
|||||
Db 173 COPMSCSGHCEVEIINNHTCNCDDVGYGPOCQ 205

RESULT 17
US-09-791-537-50403
; Sequence 50403, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50403
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-50403

Query Match          88.8%; Score 191; DB 5; Length 370;
Best Local Similarity 81.8%; Pred. No. 1.8e-13;
Matches 27; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 COPMSCSGHCEVEIINNHTCNCDDVGYGPOCQ 33
|||||
Db 160 CKPMSGSHGCVYINNHTCNCDDVGYGPECO 192

RESULT 18
US-09-791-537-132144
; Sequence 132144, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132144
```

```
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-791-537-132144

Query Match          84.2%; Score 181; DB 5; Length 376;
Best Local Similarity 84.8%; Pred. No. 2.3e-12;
Matches 28; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 COPMSCSGHCEVEIINNHTCNCDDVGYGPOCQ 33
|||||
Db 160 CHPGSCSGHCEVEIINNHTCNCDDVGYGPOCQ 192

RESULT 19
US-09-791-537-60503
; Sequence 60503, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 60503
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-60503

Query Match          78.6%; Score 169; DB 5; Length 360;
Best Local Similarity 81.8%; Pred. No. 4.7e-11;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 COPMSCSGHCEVEIINNHTCNCDDVGYGPOCQ 33
|||||
Db 160 CQPGSCNGRCEVEIINNHTCNCDDVGYGPOCQ 192

RESULT 20
US-09-791-537-37750
; Sequence 37750, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37750
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-37750

Query Match          78.6%; Score 169; DB 5; Length 372;
Best Local Similarity 81.8%; Pred. No. 4.8e-11;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 COPMSCSGHCEVEIINNHTCNCDDVGYGPOCQ 33
|||||
Db 160 CQPGSCNGRCEVEIINNHTCNCDDVGYGPOCQ 192
```

```
RESULT 21
US-09-791-537-6693
; Sequence 6693, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6693
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Rattus sp
US-09-791-537-6693
```

```
Query Match
Best Local Similarity 74.0%; Score 159; DB 5; Length 372;
Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 160 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 192
```

```
RESULT 22
US-09-791-537-81233
; Sequence 81233, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81233
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-81233
```

```
Query Match
Best Local Similarity 74.0%; Score 159; DB 5; Length 372;
Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 160 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 192
```

```
RESULT 23
US-09-791-537-118878
; Sequence 118878, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
```

```
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 118878
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-791-537-118878
```

```
Query Match
Best Local Similarity 60.9%; Score 131; DB 5; Length 611;
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 144 CPTSCGHGECVEITNNHTCKCHPGRGLRCE 176
```

```
RESULT 24
US-09-791-537-23269
; Sequence 23269, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23269
; LENGTH: 162
; TYPE: PRT
; ORGANISM: pdb 1ESL
US-09-791-537-23269
```

```
Query Match
Best Local Similarity 60.5%; Score 130; DB 5; Length 162;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 122 CPTSCGHGECVEITNNHTCKDPRFSGLRCE 154
```

```
RESULT 25
PCT-US02-23913-357
; Sequence 357, Application PC/TUS0223913
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044PC
; CURRENT APPLICATION NUMBER: PCT/US02/23913
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
```

```

; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-23913-357

```

Query Match	60.5%;	Score 130;	DB 1;	Length 610;
Best Local Similarity	63.6%;	Pred. NO. 1.5e-06;		
Matches	21;	Conservative	4;	Mismatches 8;
				Indels 0;
				Gaps 0;

Qy 1 CQPMSCSGHGECVEIINNHTCNCDDVGYYGPOCQ 33
| | | | | | | : : :
Db 143 CTNTSCSGHGEVETINNYTCKCDPFSSGLKCE 175

```

RESULT 26
US-09-791-537-22813
: Sequence 22813, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791.537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 22813
: LENGTH: 610
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-791-537-22813

```

Query Match	60.5%	Score 130;	DB 5;	Length 610;
Best Local Similarity	63.6%	Pred. No. 1.5e-06;		
Matches	21; Conservative	4; Mismatches	8; Indels	0; Gaps 0;

Qy 1 CQPWSCSGHGECEVEIINNHTCNCDDVGYGPOCQ 33
| | | | | | | | : : : :
Db 143 CTNTSCSGHGECEVEITNNYTCCKDPPGSLKCE 175

```

RESULT 27
US-09-791-537-121834
/ Sequence 121834, Application US/09791537
/ GENERAL INFORMATION:
/ APPLICANT: Bionomix, Inc.
/ APPLICANT: Debe, Derek
/ APPLICANT: Danzer, Joseph
/ TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
/ TITLE OF INVENTION: METHODS OF USE THEREOF
/ FILE REFERENCE: 261/210
/ CURRENT APPLICATION NUMBER: US/09/791,537
/ CURRENT FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 153055
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 121834
/ LENGTH: 610
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-791-537-121834

```

Query Match	60.5%	Score 130;	DB 5;	length 610;
Best Local Similarity	63.6%	Pred. No. 1.5e-06;		
Matches	21; Conservative	4; Mismatches	8; Indels	0; Gaps

QY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33
143 CTNTSCSGHGECEVETINNNTCKDCDPGFSGLKCE 175

RESULT 28
US-10-205-823-357
: Sequence 357, Application US/10205823

GENERAL INFORMATION:
 APPLICANT: Schlegel, Robert
 APPLICANT: Monahan, John E.
 APPLICANT: Endege, Wilson O.
 APPLICANT: Gannavarapu, Manjula
 APPLICANT: Gorbacheva, Bella
 APPLICANT: Hoersch, Sebastian
 APPLICANT: Kamatkar, Shubhangi
 APPLICANT: Monsey, Angela M.
 APPLICANT: Glatz, Karen

```

1  APPLICANT: Zhao, Xumei
2  APPLICANT: Anderson, Dustin
3  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
4  TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
5  TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
6  FILE REFERENCE: MRI-044
7  CURRENT APPLICATION NUMBER: US/10/205,823

```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

? NUMBER OF SEQ ID NOS: 435
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 357
? LENGTH: 610
? TYPE: PRT
? ORGANISM: Homo sapiens
? GS-10-205-823-357

```

Query Match	60.5%	Score 130;	DB 6;	Length 610;
Best Local Similarity	63.6%	Pred. No. 1.5e-06;		
Matches	21;	Conservative	4;	Mismatches 8;
				Indels 0;
				Gaps 0;

QY 1 CQPMSCSGHGECEVEIINNHTCNCDVGYGPOCQ 33
| | | | | | | | | | | | | | | | : : : :
Db 143 CTNTSCSGHGECEVEITNNYTCCKDPFGSLKCE 175

```

RESULT 29
US-09-791-537-50406
: Sequence 50406, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 50406
: LENGTH: 549
: TYPE: PRT
: ORGANISM: Rattus norvegicus
: US-09-791-537-50406

```

Query Match	59.18; Score 127; DB 5; Length 549;
-------------	-------------------------------------

; SEQ ID NO 50405
 ; LENGTH: 484
 ; TYPE: PRT
 ; ORGANISM: Sus scrofa
 US-09-791-537-50405

Query Match 58.1%; Score 125; DB 5; Length 484;
 Best Local Similarity 60.6%; Pred. No. 4,4e-06;
 Matches 20; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINNHTCNCDDVGYGPQCQ 33
 DB 144 CTPSCSGHGECEIINNSTCQCYPGFRGLQCE 176

RESULT 40
 US-09-791-537-32025
 ; Sequence 32025, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBR
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 32025
 ; LENGTH: 616
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-791-537-32025

Query Match 58.1%; Score 125; DB 5; Length 616;
 Best Local Similarity 57.6%; Pred. No. 5,4e-06;
 Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINNHTCNCDDVGYGPQCQ 33
 DB 163 CQWSCSGHGECEIINNSTCQCYPGFRGLQCE 195

Search completed: September 7, 2002, 10:22:09
 Job time: 557 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:20:37 : Search time 304.64 Seconds
(without alignments)
39.284 Million cell updates/sec

Title: US-09-119-209-2_COPY_160_193
Perfect score: 215
Sequence: 1 CQWSCSGHCEVEIINHTCNCDCVGYGQCQL 34

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/PCrus.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US081.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US080.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099.COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US100.COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US101.COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	372	15	US-09-119-209-2
2	209	97.2	382	21	US-09-760-475-2123
3	209	97.2	1078	26	US-60-212-659-523
4	205	95.3	341	21	US-09-758-449-1158
5	205	95.3	341	21	US-09-760-443-1328
6	205	95.3	372	1	PCT-US01-26675-3
7	205	95.3	385	1	PCT-US92-03970-2

8	205	95.3	385	1	PCT-US94-00909-2	Sequence 2, Appl1
9	205	95.3 <td>385</td> <td>4</td> <td>US-08-008-459-2</td> <td>Sequence 2, Appl1</td>	385	4	US-08-008-459-2	Sequence 2, Appl1
10	205	95.3 <td>385</td> <td>7</td> <td>US-08-340-539-2</td> <td>Sequence 2, Appl1</td>	385	7	US-08-340-539-2	Sequence 2, Appl1
11	205	95.3 <td>385</td> <td>8</td> <td>US-08-410-569-2</td> <td>Sequence 2, Appl1</td>	385	8	US-08-410-569-2	Sequence 2, Appl1
12	205	95.3 <td>1078</td> <td>26</td> <td>US-60-207-315-428</td> <td>Sequence 428, App</td>	1078	26	US-60-207-315-428	Sequence 428, App
13	205	95.3 <td>1078</td> <td>26</td> <td>US-60-230-435-1751</td> <td>Sequence 1751, Ap</td>	1078	26	US-60-230-435-1751	Sequence 1751, Ap
14	169	78.6	372	15	US-09-119-209-4	Sequence 4, Appl1
15	130	60.5	610	1	PCT-US99-28965-19	Sequence 19, Appl1
16	130	60.5	610	10	US-08-657-753-2	Sequence 2, Appl1
17	130	60.5	610	11	US-08-770-435-3	Sequence 3, Appl1
18	130	60.5	610	16	US-09-266-091-2	Sequence 2, Appl1
19	130	60.5	610	16	US-09-266-091A-2	Sequence 2, Appl1
20	130	60.5	610	21	US-09-784-356-122	Sequence 122, Appl1
21	130	60.5	610	22	US-09-802-640-36	Sequence 36, Appl1
22	130	60.5	610	22	US-09-857-670-19	Sequence 19, Appl1
23	130	60.5	610	24	US-10-021-660-122	Sequence 122, App
24	125	58.1	700	26	US-60-207-315-467	Sequence 467, App
25	125	58.1	830	1	PCT-US94-09395-4	Sequence 4, Appl1
26	125	58.1	830	8	US-08-449-687B-4	Sequence 4, Appl1
27	125	58.1	830	24	US-10-020-141-10	Sequence 10, Appl1
28	120	55.8	36	26	US-60-160-189-8885	Sequence 8885, Ap
29	120	55.8	37	26	US-60-160-189-10134	Sequence 10134, A
30	120	55.8	37	26	US-60-160-203-6335	Sequence 6335, Ap
31	120	55.8	38	26	US-60-160-203-5155	Sequence 5155, Ap
32	120	55.8	38	26	US-60-169-840-6794	Sequence 6794, Ap
33	120	55.8	38	26	US-60-169-867-5901	Sequence 5901, Ap
34	98	45.6	34	26	US-60-146-055-529	Sequence 529, App
35	98	45.6	173	26	US-60-139-669-865	Sequence 665, App
36	98	45.6	183	26	US-60-139-669-596	Sequence 596, App
37	98	45.6	321	16	US-09-270-767-33762	Sequence 33762, A
38	98	45.6	321	16	US-09-270-767-48979	Sequence 48979, A
39	98	45.6	321	16	US-09-270-8498-194120	Sequence 194120, A
40	98	45.6	427	26	US-60-142-896-1034	Sequence 1034, Ap
41	98	45.6	459	26	US-60-140-956-1513	Sequence 1513, Ap
42	98	45.6	496	26	US-60-145-138-711	Sequence 711, App
43	98	45.6	3396	20	US-09-614-150-19575	Sequence 19575, A
44	98	45.6	3396	26	US-60-167-217-19718	Sequence 19718, A
45	98	45.6	3396	26	US-60-191-637-19634	Sequence 19634, A

ALIGNMENTS

RESULT 1
US-09-119-209-2
Sequence 2, Application US/09119209
GENERAL INFORMATION:
APPLICANT: LASKY, LAURENCE A.
APPLICANT: STACHELL, SCOTT E.
APPLICANT: ROSEN, STEVEN D.
APPLICANT: SINGER, MARK S.
APPLICANT: YEDNOCK, TED A.
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/119,209
FILING DATE: 20-Jul-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/513278
FILING DATE: 10-AUG-1995

```

;
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-09-119-209-2
```

```

Query Match          100.0%; Score 215; DB 15; Length 372;
Best Local Similarity 100.0%; Pred. No. 4.5e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  CQPMSCSGHGECEIINNHTCNCDDVGYGPOCOL 34
Db      160 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCOL 193
```

```

RESULT 2
US-09-760-475-2123
; Sequence 2123, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2123
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-760-475-2123
```

```

Query Match          97.2%; Score 209; DB 21; Length 382;
Best Local Similarity 97.1%; Pred. No. 2.4e-16;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  CQPMSCSGHGECEIINNHTCNCDDVGYGPOCOL 34
Db      170 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCOL 203
```

```

RESULT 3
US-60-212-659-523
; Sequence 523, Application US/60212659
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00674
; CURRENT APPLICATION NUMBER: US/60/212,659
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 879
```

```

;
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
;
US-60-212-659-523
```

```

Query Match          97.2%; Score 209; DB 26; Length 1078;
Best Local Similarity 97.1%; Pred. No. 6e-16;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  CQPMSCSGHGECEIINNHTCNCDDVGYGPOCOL 34
Db      742 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCOL 775
```

```

RESULT 4
US-09-758-449-1158
; Sequence 1158, Application US/09758449
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM026
; CURRENT APPLICATION NUMBER: US/09/758,449
; CURRENT FILING DATE: 2001-01-11
; PRIORITY APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIORITY APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1478
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1158
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;
US-09-758-449-1158
```

```

Query Match          95.3%; Score 205; DB 21; Length 341;
Best Local Similarity 97.0%; Pred. No. 6.3e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  CQPMSCSGHGECEIINNHTCNCDDVGYGPOCO 33
Db      182 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCO 214
```

```

RESULT 5
US-09-760-443-1328
; Sequence 1328, Application US/09760443
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7212
; CURRENT APPLICATION NUMBER: US/09/760,443
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2164
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1328
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;
US-09-760-443-1328
```

Query Match 95.3%; Score 205; DB 21; Length 341;
Best Local Similarity 97.0%; Pred. No. 6.3e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEIINNHTCNCVDGYGPOCQ 33
|||||
DB 182 CQPMSCSGHGECEIINNHTCNCVDGYGPOCQ 214

RESULT 6
PCT-US01-26675-3
Sequence 3, Application PC/TUS0126675

GENERAL INFORMATION:
APPLICANT: Genesance Pharmaceuticals, Inc.
APPLICANT: Anastasio, Allison E
APPLICANT: Bieganski, Karyn M
APPLICANT: Kishy, Stefanie E
APPLICANT: Kumar, Anant Madan
TITLE OF INVENTION: HAPLOTYPES OF THE SELL GENE
FILE REFERENCE: SELL MWH116-PCF
CURRENT APPLICATION NUMBER: PCT/US01/26675
CURRENT FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 60/228,262
PRIOR FILING DATE: 2000-08-25
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-26675-3

Query Match 95.3%; Score 205; DB 1; Length 372;
Best Local Similarity 97.0%; Pred. No. 6.9e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEIINNHTCNCVDGYGPOCQ 33
|||||
DB 160 CQPMSCSGHGECEIINNHTCNCVDGYGPOCQ 192

RESULT 7
PCT-US92-03970-2

Sequence 2, Application PC/TUS9203970
GENERAL INFORMATION:
APPLICANT: Dana-Farber Cancer Institute, Inc.
TITLE OF INVENTION: LEUKOCYTE-ASSOCIATED CELL SURFACE
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03970
FILING DATE: 19920513
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Helne, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-152Bq9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-03970-2

Query Match 95.3%; Score 205; DB 1; Length 385;
Best Local Similarity 97.0%; Pred. No. 7.1e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEIINNHTCNCVDGYGPOCQ 33
|||||
DB 173 CQPMSCSGHGECEIINNHTCNCVDGYGPOCQ 205

RESULT 8
PCT-US94-00909-2

Sequence 2, Application PC/TUS9400909
GENERAL INFORMATION:
APPLICANT: CHIMERIC SELECTINS AS SIMULTANEOUS BLOCKING
TITLE OF INVENTION: AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00909
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/962,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/737,092
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-00909-2

Query Match 95.3%; Score 205; DB 1; Length 385;
Best Local Similarity 97.0%; Pred. No. 7.1e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPMSCGHGECVEIINHTCNCMDVGYGPOCQ 33
Db 173 CQPMSCGHGECVEIINHTCNCMDVGYGPOCQ 205

RESULT 9

US-08-008-459-2
; Sequence 2, Application US/08008459
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,459
; FILING DATE: 25-JAN-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,606
; FILING DATE: 30-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862,483
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,608
; FILING DATE: 03-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/737,092
; FILING DATE: 29-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/730,503
; FILING DATE: 08-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/700,773
; FILING DATE: 15-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/313,109
; FILING DATE: 21-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Heine, Holliday C.
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: DFCI-318XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; TELEX: 940675
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-008-459-2

Query Match 95.3%; Score 205; DB 4; Length 385;
Best Local Similarity 97.0%; Pred. No. 7,le-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPMSCGHGECVEIINHTCNCMDVGYGPOCQ 33
Db 173 CQPMSCGHGECVEIINHTCNCMDVGYGPOCQ 205

RESULT 10

US-08-340-539-2
; Sequence 2, Application US/08340539
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,539
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,459
; FILING DATE: 25-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,606
; FILING DATE: 30-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862,483
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,608
; FILING DATE: 03-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/737,092
; FILING DATE: 29-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/730,503
; FILING DATE: 08-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/700,773
; FILING DATE: 15-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/313,109
; FILING DATE: 21-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Heine, Holliday C.
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: DFCI-318XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; TELEX: 940675
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-340-539-2

Query Match 95.3%; Score 205; DB 7; Length 385;
Best Local Similarity 97.0%; Pred. No. 7,le-16;

Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 33
DB 173 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 205

RESULT 11
US-08-410-569-2

; Sequence 2, Application US/08410569
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Sperlin, Olivier G.
; TITLE OF INVENTION: LEUKOCYTE ADHESION MOLECULE-1 (LAM-1)
; TITLE OF INVENTION: AND LIGAND THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/410,569
; APPLICATION NUMBER: US/08/410,569
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,608
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: US 07/700,773
; FILING DATE: 15-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Helme, Holliday C.
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: DFCG-152EX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; TELEX: 940675
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-410-569-2

Query Match 95.3%; Score 205; DB 8; Length 385;
Best Local Similarity 97.0%; Pred. No. 7, 1e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 33
DB 173 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 205

RESULT 12
US-60-207-315-428

; Sequence 428, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00601
; CURRENT APPLICATION NUMBER: US/60/207,315

; CURRENT FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 528

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 428

; LENGTH: 1078

; TYPE: PRT

; ORGANISM: HUMAN

US-60-207-315-428

Query Match 95.3%; Score 205; DB 26; Length 1078;

Best Local Similarity 97.0%; Pred. No. 1, 8e-15;

Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 33
DB 742 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 774

RESULT 13
US-60-230-435-1751

; Sequence 1751, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1751
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1751

Query Match 95.3%; Score 205; DB 26; Length 1078;
Best Local Similarity 97.0%; Pred. No. 1, 8e-15;

Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 33
DB 742 COPWSCSGHCEVEIINNHTCNCVDGYGPGQ 774

RESULT 14
US-09-119-209-4

; Sequence 4, Application US/09119209
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: SPACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,209

```

; FILING DATE: 20-Jul-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513278
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-119-209-4
```

```

Query Match          78.6%; Score 169; DB 15; Length 372;
Best Local Similarity 81.8%; Pred. No. 1.3e-11;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 33
Db      160 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 192
```

```

RESULT 15
PCT-US99-28965-19
; Sequence 19, Application PC/TUS9928965
; GENERAL INFORMATION:
; APPLICANT: Monla, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; APPLICANT: Isis Pharmaceuticals, Inc.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
; FILE REFERENCE: ISPH-0424
; CURRENT APPLICATION NUMBER: PCT/US99/28965
; EARLIER FILING DATE: 1999-12-08
; EARLIER APPLICATION NUMBER: US 09/209,668
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US99-28965-19
```

```

Query Match          60.5%; Score 130; DB 1; Length 610;
Best Local Similarity 63.6%; Pred. No. 8.6e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 33
Db      143 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 175
```

```

RESULT 16
US-08-657-753-2
; Sequence 2, Application US/08657753
```

```

; GENERAL INFORMATION:
; APPLICANT: Klimuk, Sandra K.
; APPLICANT: Semple, Sean C.
; APPLICANT: Scherier, Peter
; APPLICANT: Hope, Michael J.
; TITLE OF INVENTION: Enhanced Efficacy of Liposomal Antisense
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/657,753
; FILING DATE: Not yet assigned
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kezer, William B.
; REGISTRATION NUMBER: 37,369
; REFERENCE/DOCKET NUMBER: 16303-003600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-657-753-2
```

```

Query Match          60.5%; Score 130; DB 10; Length 610;
Best Local Similarity 63.6%; Pred. No. 8.6e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 33
Db      143 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 175
```

```

RESULT 17
US-08-770-435-3
; Sequence 3, Application US/08770435
; GENERAL INFORMATION:
; APPLICANT: Gimbrone, Jr., Michael A.
; APPLICANT: Bevilacqua, Michael P.
; TITLE OF INVENTION: Antibodies Specific For E-selectin And The Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,435
; FILING DATE: herewith
; CLASSIFICATION: 424
```



```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  US 08/365,470
3      FILING DATE:  29-DEC-1994
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER:  US 08/102,510
6      FILING DATE:  05-AUG-1993
7      PRIOR APPLICATION DATA:
8      APPLICATION NUMBER:  US 07/850,802
9      FILING DATE:  13-MAR-1992
10     ATTORNEY/AGENT INFORMATION:
11     NAME:  Matkovicz, Karen R.
12     REGISTRATION NUMBER:  36,351
13     REFERENCE/DOCKET NUMBER:  0627,1350004
14     TELECOMMUNICATION INFORMATION:
15     TELEPHONE:  202-371-2600
16     TELEFAX:  202-371-2540
17     INFORMATION FOR SEQ ID NO: 3:
18     SEQUENCE CHARACTERISTICS:
19     LENGTH:  610 amino acids
20     TYPE:  amino acid
21     TOPOLOGY:  linear
22     US-08-770-435-3

```

Query	March	Score	130	DB	11	Length	610
		Best Local Similarity	63.6%	Pred.	0.86e-07		
Matches	21	Conservative	4	Mismatches	8	Indels	0
						Gaps	0
OY	1	COPMSCSGHGECEVEIINNHTCCMCDDVGYGPPQCQ	33				
Db	143	CTNTSCSGHGECEVEIINNHTCKCDPDSFGSLKE	175				

```

RESULT 18
US-09-266-091-2
Sequence 2, Application US/09266091
GENERAL INFORMATION:
APPLICANT: Klimuk, Sandra K
APPLICANT: Semple, Sean C
APPLICANT: Scherrier, Peter
APPLICANT: Hope, Michael J
APPLICANT: Bennett, C. Frank
TITLE OF INVENTION: ENHANCED EFFICACY OF LIPOSOMAL ANTISENSE THERAPY
FILE REFERENCE: ISPH-0342
CURRENT APPLICATION NUMBER: US/09/266,091
CURRENT FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 08/657,753
PRIOR FILING DATE: 1996-05-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 610
TYPE: PRT
ORGANISM: Homo sapiens
US-09-266-091-2

```

Query Match	60.5%	Score 130	DB 16	Length 610
Best Local Similarity	63.6%	Pred. No. 8	6e-07	
Matches 21	Conservative 4	Mismatches 8	Indels 0	Gaps 0

QY	1	CGPMSGSGHGECEVETINNHTCNCDDVGYGGPOCQ	33
Db	143	CTNTSCSGHGECEVETINNHTCNCDDVGYGGPOCQ	175

RESULT 19
US-09-266-091A-2
; Sequence 2, Application US/09266091A
; GENERAL INFORMATION:
; APPLICANT: Klimuk, Sandra K.
; Semple, Sean C.
; Schierter, Peter

```

1 Hope, Michael J.
2 TITLE OF INVENTION: Enhanced Efficacy of Liposomal Antisense
3 NUMBER OF SEQUENCES: 10
4 CORRESPONDENCE ADDRESS:
5 ADDRESSEE: Townsend and Townsend and Crew LLP
6 STREET: Two Embarcadero Center, Eighth Floor
7 CITY: San Francisco
8 STATE: California
9 COUNTRY: USA
10 ZIP: 94111-3834
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: PatentIn Release #1.0, Version #1.30
17
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/09/266,091A
20 FILING DATE: 10-Mar-1999
21
22 CLASSIFICATION: <Unknown>
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: 08/657,753
26 FILING DATE: <Unknown>
27
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Kezer, William B.
30 REGISTRATION NUMBER: 37,369
31 REFERENCE/DOCKET NUMBER: 16303-003600
32
33 TELECOMMUNICATION INFORMATION:
34 TELEPHONE: (415) 576-0200
35 TELEFAX: (415) 576-0300
36
37 INFORMATION FOR SEQ ID NO: 2:
38
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 610 amino acids
41 TYPE: amino acid
42 TOPOLOGY: linear
43
44 MOLECULE TYPE: protein
45
46 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
47
48 US-09-266-091A-2

```

Query Match	60.5%	Score 130	DB 16	length 610
Best Local Similarity	63.6%	Pred. 0.86e-07		
Matches	21	Conservative	4	Mismatches 8
				Indels 0
				Gaps 0

QY	1	COPWSCSGHGCEVETLNNHTCMCDVGYGGPOCQ	33
			: :
Db	143	CTNTSCSGHGCEVETLNNHTCMCDGFGSLKCE	175

```

RESULT 20
US-09-784-356-122
; Sequence 122, Application US/09784356
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynnne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/09/784,356
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 60/7148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-784-356-122

```


APPLICANT: Board of Regents of the University of Oklahoma
TITLE OF INVENTION: Expression Control Sequences of the P-Selectin Gene
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09395
FILING DATE: 19-AUG-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-09395-4

Query Match 58.1%; Score 125; DB 1; Length 830;
Best Local Similarity 57.6%; Pred. No. 4.5e-06;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CQWSCSGHCEVEIINHTCNCVDVGYGPQOC 33
Db 163 CODMSCSKQGCLETIGNYTCSCYPFGPCE 195

RESULT 26
US-08-449-687B-4
Sequence 4, Application US/08449687B
GENERAL INFORMATION:
APPLICANT: McEver, Rodger P.
APPLICANT: Pan, Junliang
TITLE OF INVENTION: Expression Control Sequences of the
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,687B
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,158
FILING DATE: 20-AUG-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/320,408
FILING DATE: 08-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-449-687B-4

Query Match 58.1%; Score 125; DB 8; Length 830;
Best Local Similarity 57.6%; Pred. No. 4.5e-06;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CQWSCSGHCEVEIINHTCNCVDVGYGPQOC 33
Db 163 CODMSCSKQGCLETIGNYTCSCYPFGPCE 195

RESULT 27
US-10-020-141-10
Sequence 10, Application US/10020141
GENERAL INFORMATION:
APPLICANT: McCarthy, Jeanette
APPLICANT: Ableson, Allen
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
FILE REFERENCE: MMI-002
CURRENT APPLICATION NUMBER: US/10/020,141
CURRENT FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/313,097
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: US 60/327,485
PRIOR FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 830
TYPE: PRT
ORGANISM: Homo sapiens
US-10-020-141-10

Query Match 58.1%; Score 125; DB 24; Length 830;
Best Local Similarity 57.6%; Pred. No. 4.5e-06;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CQWSCSGHCEVEIINHTCNCVDVGYGPQOC 33
Db 163 CODMSCSKQGCLETIGNYTCSCYPFGPCE 195

RESULT 28
US-60-160-189-8885
Sequence 8885, Application US/60160189
GENERAL INFORMATION:
APPLICANT: BONA221, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
FILE REFERENCE: CLO00112
CURRENT APPLICATION NUMBER: US/60/160,189
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 10162
SOFTWARE: FastSeq for Windows Version 4.0


```

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

```

```
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48979
; LENGTH: 321
; TYPE: PRF
; ORGANISM: Drosophila melanogaster
US-09-270-767-48979
```

```
Query Match                      45.6%; Score 98; DB 16; Length 321;
Best Local Similarity            39.4%; Pred. No. 0.003;
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCSGHGEVCEIINNHTCNCVDVGYGPQCQ 33
      1: 1: |||::: ||| 11
Db       138 CENRPCANEGTCVDLVGYSCNCEPGYTGKNCQ 170
```

```
RESULT 39
US-09-270-849B-194120
; Sequence 194120, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 194120
; LENGTH: 321
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-194120
```

```
Query Match                      45.6%; Score 98; DB 16; Length 321;
Best Local Similarity            39.4%; Pred. No. 0.003;
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCSGHGEVCEIINNHTCNCVDVGYGPQCQ 33
      1: 1: |||::: ||| 11
Db       138 CENRPCANEGTCVDLVGYSCNCEPGYTGKNCQ 170
```

```
RESULT 40
US-60-142-896-1034
; Sequence 1034, Application US/60142896
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000052
; CURRENT APPLICATION NUMBER: US/60/142,896
; CURRENT FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 1584
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1034
; LENGTH: 427
; TYPE: PRF
; ORGANISM: Drosophila
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(427)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-142-896-1034
```

```
Query Match                      45.6%; Score 98; DB 26; Length 427;
Best Local Similarity            39.4%; Pred. No. 0.0039;
```

```
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY      1 CQPMSCSGHGEVCEIINNHTCNCVDVGYGPQCQ 33
      1: 1: |||::: ||| 11
Db       72 CENRPCANEGTCVDLVGYSCNCEPGYTGKNCQ 104
```

Search completed: September 7, 2002, 10:20:38
Job time: 486 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:42 : Search time 28.63 Seconds
(without alignments)
29.007 Million cell updates/sec

Title: US-09-119-209-2_COPY_160_193
Perfect score: 215
Sequence: 1 COPWCSGHGECVEIINHHHCNDGVYGPQCL 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	372	2	US-08-513-278-2
2	215	100.0	372	1	Sequence 2, Appl
3	205	95.3	385	1	Patent No. 5514582
4	205	95.3	385	2	Sequence 2, Appl
5	191	88.8	36	1	US-08-340-539A-22
6	169	78.6	34	6	5514582-21
7	169	78.6	36	1	US-08-340-539A-20
8	169	78.6	372	2	US-08-513-278-4
9	169	78.6	372	6	5514582-4
10	159	74.0	36	1	US-08-340-539A-21
11	130	60.5	36	1	US-08-340-539A-23
12	130	60.5	610	1	US-08-365-470-3
13	130	60.5	610	3	US-08-209-668-19
14	130	60.5	610	4	US-09-009-490A-89
15	130	60.5	610	6	5217870-2
16	127	59.1	36	1	US-08-340-539A-28
17	126	58.6	36	1	US-08-340-539A-24
18	125	58.1	36	1	US-08-340-539A-26
19	125	58.1	484	2	US-08-252-493C-9
20	125	58.1	484	3	US-09-276-197-9
21	125	58.1	830	4	US-08-110-158-4
22	125	58.1	830	5	PCT-US91-05059-2
23	125	58.1	830	6	5378464-2
24	123	57.2	36	1	US-08-340-539A-25
25	120	55.8	36	1	US-08-340-539A-27
26	93	43.3	520	4	US-09-068-740A-3
27	93	43.3	642	3	US-08-872-855-10

28 93 43.3 702 4 US-09-068-740A-4
29 93 43.3 723 1 US-09-068-740A-9
30 93 43.3 2523 4 US-08-185-432-18
31 90 41.9 721 3 US-08-872-855-7
32 90 41.9 721 4 US-08-981-392-5
33 89 41.4 2471 1 US-08-185-432-16
34 89 41.4 2471 1 US-08-083-590A-19
35 89 41.4 2471 3 US-08-533-384-19
36 88 40.9 713 3 US-08-872-855-5
37 87 40.5 728 4 US-08-981-392-2
38 87 40.5 729 4 US-08-872-855-8
39 86 40.0 34 6 5514582-22
40 86 40.0 2703 1 US-08-185-432-19
41 84 39.1 717 3 US-08-872-855-9
42 84 39.1 1523 4 US-09-182-024A-2
43 84 39.1 2199 5 PCT-US95-11684-2
44 83 38.6 1139 1 US-08-537-210A-4
45 83 38.6 1139 4 US-09-113-825-4

ALIGNMENTS

RESULT 1
US-08-513-278-2
Sequence 2, Application US/08513278
Patent No. 5840844
GENERAL INFORMATION:
APPLICANT: LASKY, LAURENCE A.
APPLICANT: STACHELL, SCOTT E.
APPLICANT: ROSEN, STEVEN D.
APPLICANT: SINGER, MARK S.
APPLICANT: YEDNICK, TED A.
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,278
FILING DATE: 10-AUG-1995
CLASSIFICATION: 5530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/059027
FILING DATE: 06-MAY-1993
APPLICATION NUMBER: 07/786149
FILING DATE: 31-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/315015
FILING DATE: 23-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Deger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 565D1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/952-9216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-513-278-2

Sequence 4, Appl
Sequence 9, Appl
Sequence 18, Appl
Sequence 7, Appl
Sequence 5, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 5, Appl
Sequence 2, Appl
Sequence 19, Appl
Patent No. 5514582
Sequence 19, Appl
Sequence 9, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl

Query Match 100.0%; Score 215; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 5,6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 34
Db 160 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 193

RESULT 2
5514582-2

; Patent No. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
; IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO: 2:
; LENGTH: 372
5514582-2

Query Match 100.0%; Score 215; DB 6; Length 372;
Best Local Similarity 100.0%; Pred. No. 5,6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 34
Db 160 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 193

RESULT 3

; Sequence 2, Application US/08340539A
; Patent No. 5808025
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,539A
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,459
; FILING DATE: 25-JAN-1993

; ATTORNEY/AGENT INFORMATION:
; NAME: Gunnison, Jane
; REGISTRATION NUMBER: 38,479
; REFERENCE/DOCKET NUMBER: CG-104 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-340-539A-2

Query Match 95.3%; Score 205; DB 1; Length 385;
Best Local Similarity 97.0%; Pred. No. 7,6e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINNHTCNCNVGYGPOCQ 33
Db 173 COPWSCSGHGEVEIINNHTCNCNVGYGPOCQ 205

RESULT 4

; Sequence 2, Application US/08461592B
; Patent No. 5834425
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,592B
; FILING DATE:
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/340,539
; FILING DATE: 16-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,459
; FILING DATE: 25-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CG-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELETYPE: 14-8367
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-592B-2

Query Match 95.3%; Score 205; DB 2; Length 385;
Best Local Similarity 97.0%; Pred. No. 7, 6e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 COPWSCSGHCEVEIINHTCNCDDVGYGPQOQ 33
Db 173 COPWSCSGHCEVEIINHTCNCDDVGYGPQOQ 205

RESULT 5
US-08-340-539A-22
Sequence 22, Application US/08340539A
Patent No. 5808025

GENERAL INFORMATION:

APPLICANT: Tedder, Thomas F.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Gunnison, Jane

REGISTRATION NUMBER: 38,479

REFERENCE/DOCKET NUMBER: CG-104 CON

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-340-539A-22

Query Match 88.8%; Score 191; DB 1; Length 36;
Best Local Similarity 81.8%; Pred. No. 2, 7e-15;
Matches 27; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 COPWSCSGHCEVEIINHTCNCDDVGYGPQOQ 33
Db 3 CKPWSCSGHGCEVEIINHTCNCDDVGYGPQOQ 35

RESULT 6
5514582-21
Patent No. 5514582

APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.

TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID

NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 986,931

FILING DATE: 08-DEC-1992

APPLICATION NUMBER: 808,122

FILING DATE: 16-DEC-1991

APPLICATION NUMBER: 440,625

FILING DATE: 22-NOV-1989

APPLICATION NUMBER: 315,015

FILING DATE: 23-FEB-1989

SEQ ID NO: 21;
LENGTH: 34

5514582-21

Query Match 78.6%; Score 169; DB 6; Length 34;
Best Local Similarity 81.8%; Pred. No. 7, 5e-13;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 COPWSCSGHCEVEIINHTCNCDDVGYGPQOQ 33
Db 1 COPWSCSGHCEVEIINHTCNCDDVGYGPQOQ 33

RESULT 7
US-08-340-539A-20

Sequence 20, Application US/08340539A
Patent No. 5808025

GENERAL INFORMATION:

APPLICANT: Tedder, Thomas F.

TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS

TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESS: FISH & NEAVE

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30.
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Gunnison, Jane

REGISTRATION NUMBER: 38,479

REFERENCE/DOCKET NUMBER: CG-104 CON

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-340-539A-20

Query Match 78.6%; Score 169; DB 1; Length 36;
Best Local Similarity 81.8%; Pred. No. 7, 9e-13;

Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINNHTCNCVGYGPOCQ 33
||| ||:| ||||| ||||| || ||||| ||
Db 3 CQWSCNGRGECEIINNHTCICDAGYGPQCQ 35

RESULT 8
US-08-513-278-4
; Sequence 4, Application US/08513278
; Patent No. 5840844
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNICK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; City: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: palin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513, 278
; FILING DATE: 10-AUG-1995
; CLASSIFICATION: 5530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 06-MAY-1993
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 56501C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-513-278-4

Query Match 78.6%; Score 169; DB 2; Length 372;
Best Local Similarity 81.8%; Pred. No. 7.9e-12;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINNHTCNCVGYGPOCQ 33
||| ||:| ||||| ||||| || ||||| ||
Db 160 CQWSCNGRGECEIINNHTCICDAGYGPQCQ 192

RESULT 9
5514582-4
; Patent No. 5514582
; APPLICANT: CARON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID

; IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO: 4:
; LENGTH: 372
5514582-4

Query Match 78.6%; Score 169; DB 6; Length 372;
Best Local Similarity 81.8%; Pred. No. 7.9e-12;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINNHTCNCVGYGPOCQ 33
||| ||:| ||||| ||||| || ||||| ||
Db 160 CQWSCNGRGECEIINNHTCICDAGYGPQCQ 192

RESULT 10
US-08-340-539A-21
; Sequence 21, Application US/08340539A
; Patent No. 5808025
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; City: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,539A
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,459
; FILING DATE: 25-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gunnison, Jane
; REGISTRATION NUMBER: 38,479
; REFERENCE/DOCKET NUMBER: CG-104 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-340-539A-21

Query Match 74.0%; Score 159; DB 1; Length 36;
Best Local Similarity 78.6%; Pred. No. 1e-11;
Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 33
Db 3 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 35

RESULT 11

US-08-340-539A-23
; Sequence 23, Application US/08340539A
; Patent No. 5808025
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULANEUS
; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340.539A
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,459
; FILING DATE: 25-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gunnison, Jane
; REGISTRATION NUMBER: 38,479
; REFERENCE/DOCKET NUMBER: CG-104 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-340-539A-23

Query Match 60.5%; Score 130; DB 1; Length 36;
Best Local Similarity 63.6%; Pred. No. 1.8e-08;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 33
Db 3 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 35

RESULT 12

US-08-365-470-3
; Sequence 3, Application US/08365470
; Patent No. 5632991
; GENERAL INFORMATION:
; APPLICANT: Gimprone, Jr., Michael A.
; TITLE OF INVENTION: Antibodies Specific For E-selectin And The Uses
; TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,470
; FILING DATE: herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,510
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/850,802
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Markowicz, Karen R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 0627,1350003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2540
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-365-470-3

Query Match 60.5%; Score 130; DB 1; Length 610;
Best Local Similarity 63.6%; Pred. No. 3e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 33
Db 143 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 175

RESULT 13

US-09-209-668-19
; Sequence 19, Application US/09209668A
; Patent No. 614517
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; TITLE OF INVENTION: ALPHA-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
; FILE REFERENCE: ISPH-0336
; CURRENT APPLICATION NUMBER: US/09/209,668A
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-209-668-19

Query Match 60.5%; Score 130; DB 3; Length 610;
Best Local Similarity 63.6%; Pred. No. 3e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CQPMSCGSGHCEVEIINNHTCNCVGYGPOCQ 33

Db 143 CTNTSCGHEGCEVTINNTCKDCDPGFSGLKCE 175

RESULT 14

US-09-009-490A-89
; Sequence 89, Application US/09009490A
; Patent No. 6300491
; GENERAL INFORMATION:
; APPLICANT: Bennett and Mirabelli
; TITLE OF INVENTION: Oligonucleotide Modulation
; TITLE OF INVENTION: of Cell Adhesion
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Office of Jane Massey Licata
; STREET: 66 East Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WORDPERFECT 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,490A
; FILING DATE: January 20, 1998
; CLASSIFICATION: 514
; APPLICATION DATA:
; APPLICATION NUMBER: 440,740
; FILING DATE: May 12, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 063,167
; FILING DATE: May 17, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 969,151
; FILING DATE: February 10, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 007,997
; FILING DATE: January 20, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 939,855
; FILING DATE: September 2, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 567,286
; FILING DATE: August 14, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 810-1515
; TELEFAX: (609) 810-1454
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: no
US-09-009-490A-89

Query Match 60.5%; Score 130; DB 4; Length 610;
Best Local Similarity 63.6%; Pred. No. 3e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 COPWSCGHEGCEVTINNTCKDCDPGFSGLKCE 33
Db 143 CTNTSCGHEGCEVTINNTCKDCDPGFSGLKCE 175

RESULT 15
5217870-2
; Patent No. 5217870
; APPLICANT: HESSTON, CATHERINE A.; LOBB, ROY R.; GOELZ, SUSAN E.
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AGAINST CDX
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/345,151
; FILING DATE: 28-APR-1989
; SEQ ID NO: 2
; LENGTH: 610
5217870-2

Query Match 60.5%; Score 130; DB 6; Length 610;
Best Local Similarity 63.6%; Pred. No. 3e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 COPWSCGHEGCEVTINNTCKDCDPGFSGLKCE 33
Db 143 CTNTSCGHEGCEVTINNTCKDCDPGFSGLKCE 175

RESULT 16
US-08-340-539A-28
; Sequence 28, Application US/08340539A
; Patent No. 5808025
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,539A
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,459
; FILING DATE: 25-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gunnison, Jane
; REGISTRATION NUMBER: 38,479
; REFERENCE/DOCKET NUMBER: CG-104 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-340-539A-28

Query Match 59.1%; Score 127; DB 1; Length 36;
Best Local Similarity 57.6%; Pred. No. 4e-08;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

```

OY      1 CQPMSCSGHGECEVEIINNHTCNCDDVGYGPOCQ 33
        || ||| ||| : | : ||| : | : ||| : | :
DB      3 CQDMSCSKQGECEIETIGNTCSQYPGFYGPCE 35

```

```

US-08-340-539A-24
SEQUENCE 24, APPLICATION US/08340539A
PATENT NO. 5608025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHEMICR SELECTINS AS SIMULTANEOUS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO.: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-24

```

Query Match	58.6%	Score 126	DB 1	Length 36
Best Local Similarity	57.6%	Pred. 1.6	5.2e-09	
Matches 19	Conservative 5	Mismatches 9	Indels 0	Gaps 0
QY	1	CQPMSCSGHGECVEIINNHCTNCNDGYGGPOCQ	33	
ob	3	CTNASCSCGHGCIETINSTYCTCHGPGFGPNE	35	

RESULT 18
US-08-340-539A-26
Sequence 26, Application US/08340539A
Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas

```

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: Linear

MOLECULE TYPE: protein
OS-08-340-539A-26

```

Query Match	58.1%	Score 125	DB 1:	Length 36;
Best Local Similarity	57.6%	Pred. No. 6.7e-08;		
Matches	19;	Conservative	6;	Mismatches 8;
				Indels 0;
				Gaps 0;
OY	1	COPWSCSGHGCVEELIINNHCTCNCVDGYGTPOCQ	33	
Db	3	CQDMSCSKOGECLFETIGNVTGTCSCYGFPGPEEC	35	

RESULT 19
US-08-252-493C-9
Sequence 9, Application US/08252493C
Patent No. 5891645
GENERAL INFORMATION:
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Evans, Mark J.
APPLICANT: Malis, Louis A.
TITLE OF INVENTION: PORCINE E-SELECTIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park, Box 15
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 kb storage
COMPUTER: PC compatible
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,493C
FILING DATE: June 1, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

DESCRIPTION: porcine E-selectin

APPLICANT: Oklahoma

TITLE OF INVENTION: Peptides
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street, Suite 3100
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05059
FILING DATE: 19910717
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/320408
FILING DATE: 08-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554199
FILING DATE: 17-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRFL10CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6558
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: YES
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Blood
CELL TYPE: Endothelial
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 4..25
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 60..158
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 131..150
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 163..174
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 168..183
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 185..194
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 200..244
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 213..226
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 230..257
NAME/KEY: Disulfide-bond
LOCATION: 262..306
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 275..288
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 292..319
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 324..368
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 337..350
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 354..381
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 386..430
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 399..412
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 416..443
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 448..492
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 461..474
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 478..505
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 510..554
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 523..536
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 540..567
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 572..616
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 585..616
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 602..629
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 642..686
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 655..668
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 672..699
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 704..748
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 717..730
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 734..761
FEATURE:
NAME/KEY: Binding-site

TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-27

Query Match 55.8%; Score 120; DB 1; Length 36;
Best Local Similarity 51.5%; Pred. No. 2,4e-07;
Matches 17; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHHTCNCDDVGYGPOCO 33
DB 3 CODMSCNMGKCIETIGSYTSCYGFYGPCE 35

RESULT 26
US-09-068-740A-3
Sequence 3, Application US/09068740A
GENERAL INFORMATION:
APPLICANT: SAKANO, SEIJI
TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
FILE REFERENCE: KP-8447
CURRENT APPLICATION NUMBER: US/09/068,740A
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: JP 7-299611
PRIOR FILING DATE: 1995-11-17
PRIOR APPLICATION NUMBER: JP 7-311811
PRIOR FILING DATE: 1995-11-30
PRIOR APPLICATION NUMBER: PCT/JP96/03356
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 48
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 520
TYPE: PRT
ORGANISM: Homo sapiens
US-09-068-740A-3

Query Match 43.3%; Score 93; DB 4; Length 520;
Best Local Similarity 38.2%; Pred. No. 0.0036;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHHTCNCDDVGYGPOCOL 34
DB 311 CDPSPCKNGSCCTDIENSYSCYCPGFYGCICEL 344

RESULT 27
US-08-872-855-10
Sequence 10, Application US/08872855
Patent No. 6121045
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean
APPLICANT: Gearity, David
TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,855
FILING DATE: 11-JUN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MAA-003.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 642 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-872-855-10

Query Match 43.3%; Score 93; DB 3; Length 642;
Best Local Similarity 41.2%; Pred. No. 0.0044;
Matches 14; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHHTCNCDDVGYGPOCOL 34
DB 428 CASSPCANGCICVDVAVNSYTCSTLGYGKDCCTL 461

RESULT 28
US-09-068-740A-4
Sequence 4, Application US/09068740A
Patent No. 6337387
GENERAL INFORMATION:
APPLICANT: SAKANO, SEIJI
TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
FILE REFERENCE: KP-8447
CURRENT APPLICATION NUMBER: US/09/068,740A
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: JP 7-299611

;; PRIOR FILING DATE: 1995-11-17
;; PRIOR APPLICATION NUMBER: JP 7-311811
;; PRIOR FILING DATE: 1995-11-30
;; PRIOR APPLICATION NUMBER: PCT/JP96/03356
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 48
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO: 4
;; LENGTH: 702
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-068-740A-4

Query Match 43.3%; Score 93; DB 4; Length 702;
Best Local Similarity 38.2%; Pred. No. 0.0048;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHEGCEVEIINNHTCNCVGYGPOCQL 34
DB 311 CDPSPCKNGGSCDTLENSYSCTCPGFYKICEL 344

RESULT 29
US-09-068-740A-9
; Sequence 9, Application US/09068740A
; Patent No. 6337387
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; APPLICANT: ITOH, AKIRA
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8447
; CURRENT APPLICATION NUMBER: US/09/068,740A
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: JP 7-299611
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: JP 7-311811
; PRIOR FILING DATE: 1995-11-30
; PRIOR APPLICATION NUMBER: PCT/JP96/03356
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO: 9
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-068-740A-9

Query Match 43.3%; Score 93; DB 4; Length 723;
Best Local Similarity 38.2%; Pred. No. 0.0049;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHEGCEVEIINNHTCNCVGYGPOCQL 34
DB 332 CDPSPCKNGGSCDTLENSYSCTCPGFYKICEL 365

RESULT 30
US-08-185-432-18
; Sequence 18, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Bussseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS

;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/185,432
;; FILING DATE: 21-JAN-1994
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mirock, S. Leslie
;; REGISTRATION NUMBER: 18,872
;; REFERENCE/DOCKET NUMBER: 7326-006
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2523 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
US-08-185-432-18

Query Match 43.3%; Score 93; DB 1; Length 2523;
Best Local Similarity 45.5%; Pred. No. 0.017;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHEGCEVEIINNHTCNCVGYGPOCQ 33
DB 910 COPNCHNGGSCDSJNMFPCNCPAGFRGPKE 942

RESULT 31
US-08-872-855-7
; Sequence 7, Application US/08872855
; Patent No. 6121045
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean
; APPLICANT: Gearing, David
; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/872,855
;; FILING DATE: 11-JUN-1997
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Arnold, Beth E.
;; REGISTRATION NUMBER: 35,430
;; REFERENCE/DOCKET NUMBER: MAA-003.02
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-832-1000
;; TELEFAX: 617-832-7000

;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 721 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: 1linear
;; MOLECULE TYPE: protein
US-08-872-855-7

Query Match 41.9%; Score 90; DB 3; Length 721;
Best Local Similarity 38.2%; Pred. No. 0.011;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPWSCSGHGECEIINNHTCNCDDVGYGPOCOL 34
Db 334 CDANPCKNKGSCSDLENSYTCSCPPGFYGNCEL 367

RESULT 32
US-08-981-392-5
; Sequence 5, Application US/08981392
; Patent No. 6262025
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos Manuel Pinto
; APPLICANT: Lewis, Julian Hart
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Gray, Grace
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES
; TITLE OF INVENTION: OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,392
; FILING DATE: 22-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Anliet, Adriane M.
; REGISTRATION NUMBER: 32,605
; REFERENCE/DOCKET NUMBER: 7326-038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 721 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-981-392-5

Query Match 41.9%; Score 90; DB 4; Length 721;
Best Local Similarity 38.2%; Pred. No. 0.011;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPWSCSGHGECEIINNHTCNCDDVGYGPOCOL 34
Db 334 CDANPCKNKGSCSDLENSYTCSCPPGFYGNCEL 367

RESULT 33
US-08-185-432-16
; Sequence 16, Application US/08185432
; Patent No. 5750652

;; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DETEKT PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2471 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-185-432-16

Query Match 41.4%; Score 89; DB 1; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.046;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPWSCSGHGECEIINNHTCNCDDVGYGPOCOL 34
Db 498 COSNPCVNNGCVDKVNRFQCLCPGFTGPVQOI 531

RESULT 34
US-08-083-590A-19
; Sequence 19, Application US/08083590A
; Patent No. 5786158
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2471 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-083-590A-19

Query Match 41.4%; Score 89; DB 1; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.046;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVETIINNHTCNCVGYGPOCOL 34
Db 498 CQSNPCVNGCVDKXNRFQCLCPGFTGYVCOI 531

RESULT 35
US-08-532-384-19
Sequence 19, Application US/08512384
Patent No. 6083904
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2471 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-532-384-19

Query Match 41.4%; Score 89; DB 3; Length 2471;
Best Local Similarity 35.3%; Pred. No. 0.018;
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVETIINNHTCNCVGYGPOCOL 34
Db 498 CQSNPCVNGCVDKXNRFQCLCPGFTGYVCOI 531

RESULT 36
US-08-872-855-5
Sequence 5, Application US/0872855
Patent No. 6121045
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean
TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,855
FILING DATE: 11-JUN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MAA-003.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-7000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-872-855-5

Query Match 40.9%; Score 88; DB 3; Length 713;
Best Local Similarity 35.3%; Pred. No. 0.018;
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVETIINNHTCNCVGYGPOCOL 34
Db 330 CAPSPCKNGSGCTLDEYSCTCPGFTGYKVCCL 363

RESULT 37

APPLICANT: Buseau, Isabelle
APPLICANT: Diedrich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2703 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-185-432-19

Query Match 40.08; Score 86; DB 1; Length 2703;
Best Local Similarity 42.48; Pred. No. 0.11;
Matches 14; Conservative 4; Mismatches 15; Indels 0; Gaps 0;
QY 1 CQPMSCGSHGECVETINNHTCNCDVGYGPQCQ 33
Db 1026 CTSSCLNGSCIDGNGINCSCLAGISGANCO 1058

Search completed: September 7, 2002, 10:14:43
Job time: 276 sec

1	215	100.0	369	17	AA898135	Human	lymphocyte c
2	215	100.0	370	17	AA898136	Human	lymphocyte c
3	215	100.0	370	17	AA898138	Human	lymphocyte c
4	215	100.0	370	17	AA898130	Human	lymphocyte c
5	215	100.0	371	17	AA898109	Human	lymphocyte c
6	215	100.0	371	17	AA898110	Human	lymphocyte c
7	215	100.0	371	17	AA898111	Human	lymphocyte c
8	215	100.0	371	17	AA898112	Human	lymphocyte c
9	215	100.0	371	17	AA898113	Human	lymphocyte c
10	215	100.0	371	17	AA898117	Human	lymphocyte c
11	215	100.0	371	17	AA898118	Human	lymphocyte c

12	215	100.0	371	17	AA9898119	Human lymphocyte c
13	215	100.0	371	17	AA9898120	Human lymphocyte c
14	215	100.0	371	17	AA9898121	Human lymphocyte c
15	215	100.0	371	17	AA9898122	Human lymphocyte c
16	215	100.0	371	17	AA9898123	Human lymphocyte c
17	215	100.0	371	17	AA9898124	Human lymphocyte c
18	215	100.0	371	17	AA9898129	Human lymphocyte c
19	215	100.0	372	12	AA912469	Human lymphocyte ce
20	215	100.0	372	13	AA922802	Human lymphocyte h
21	215	100.0	372	13	AA924026	Sequence of human
22	215	100.0	372	14	AA937960	Human lymphocyte H
23	215	100.0	372	14	AA938908	HuLHR. Homo sapie
24	215	100.0	372	16	AA976506	Human LHR. Homo s
25	215	100.0	372	16	AA983050	Human LHR. Homo s
26	215	100.0	372	17	AA9898106	Human lymphocyte c
27	215	100.0	372	17	AA9898133	Human lymphocyte c
28	215	100.0	372	19	AAW37781	Homo sapiens Lymph
29	215	100.0	372	20	AAW73264	Human lymphocyte h
30	215	100.0	374	17	AA9898133	Human lymphocyte c
31	215	100.0	374	17	AA9898132	Human lymphocyte c
32	215	100.0	374	17	AA9898134	Human lymphocyte c
33	215	100.0	374	17	AA9898135	Human lymphocyte c
34	213	99.1	371	17	AA9898115	Human lymphocyte c
35	209	97.2	371	17	AA9898116	Human lymphocyte c
36	207	96.3	371	17	AA9898114	Human lymphocyte c
37	205	95.3	196	21	AA957762	Human L-selectin a
38	205	95.3	363	13	AA922551	T lymphocyte-speci
39	205	95.3	363	17	AA991443	Human leu8 antigen
40	205	95.3	363	21	AA996184	Human T-cell speci
41	205	95.3	372	22	AA968334	Amino acid sequenc
42	205	95.3	384	11	AA905645	Lymphocyte associa
43	205	95.3	385	11	AA920815	T lymphocyte-speci
44	205	95.3	385	13	AA928753	Human LAM-1. Homo
45	205	95.3	385	14	AA931042	LAM-1 encoded by B

ALIGNMENTS	
RESULT	1
ID	AAR98125
AA	AAR98125 standard; Protein; 369 AA.
XX	
AC	AAR98125;
XX	
DT	01-NOV-1996 (first entry)
DE	
XX	Human lymphocyte cell surface glycoprotein (HLHR) variant.
XX	
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW	immunomodulator; cell adhesion; graft rejection; inflammation;
KM	metastasis.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Region
FT	Location/Qualifiers
FT	1..37
FT	/label= Signal region.
FT	38..152
FT	/label= Lectin domain.
FT	157..190
FT	/label= EGF domain.
FT	194..255
FT	/label= Complement binding repeat 1.
FT	256..314
FT	/label= Complement binding repeat 2.
FT	330..352
FT	/label= Transmembrane domain.
FT	353..369
FT	/label= Cytoplasmic domain.
XX	
PN	USS514582-A.

XX 07-MAY-1996.
PD
XX
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX
PA (GETH) GENENTECH INC.
XX
XX
PI Capon DJ, Lasky LA;
XX
XX WPI; 1996-238773/24.
DR
XX
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX
PS Disclosure; Page 19; 41pp; English.

A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region, fused at its C-terminus with the N-terminus of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targeting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the *in vitro* assay of LBP and their targets, or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particular cells. Typical applications are as antiviral, neuromodulating and immunomodulating agents, or as modulators of cell adhesion (e.g. in treating graft rejection; inflammation; metastasis of lymphoma etc.). The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR98106 are given in AAR98107-989135. This variant is a Glyc6-Ile97 deletion mutant.

SQ Sequence 369 AA;

Query Match	100.0%;	Score 215;	DB 17;	Length 369;
Best Local Similarity	100.0%;	Pred. No. 3e-14;		
Matches 34;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 CQPMSCSGHGECEIINNHNTCNDVGYYGPOQL 34
 |||||
 Db 157 CQPMSCSGHGECEIINNHNTCNDVGYYGPPQQL 190

RESULT	2
AAR98126	
ID	AAR98126 standard; Protein; 370 AA.

AC AAR98126;

DT 01-NOV-1996 (first entry)

Human lymphocyte cell surface glycoprotein (HULHR) variant.

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.

05 Homo sapiens.

xxi

FT	Key	Location/Qualifiers
FT	Region	1..37
FT	/label=	Signal region.
FT	Domain	38..153
FT	/label=	Lectin domain.
FT	Domain	158..191
FT	/label=	EGF domain.
FT	Binding-site	195..256
FT	/label=	Complement binding repeat 1.
FT	Binding-site	257..315
FT	/label=	Complement binding repeat 2.
FT	Domain	331..353
FT	/label=	Transmembrane domain.
FT	Domain	354..370
FT	/label=	Cytoplasmic domain.

PN	US5514582-A,	
XX		
PD	07-MAY-1996,	
XX		
PF	23-FEB-1989;	89US-0315015
XA		
FR	22-NOV-1989;	89US-0440625
FR	23-FEB-1989;	89US-0315015
PR	16-DEC-1991;	91US-08080122
PR	08-DEC-1992;	92US-09869311
PR	28-JAN-1994;	92US-01855570
PR	08-DEC-1992;	94US-00868570

PA (GETH) GENENTECH INC.

PI Capon DJ, Lasky LA;

DR WPI; 1996-238773/24.

PT Nucleic acid encoding

PT diagnosis and treatment

PS Disclosure; Page 19;

CC A hybrid immunoglobulin

CC region; fused at its

CC immunoglobulin super

CC the adhesion/targett

more than one ligand

CC LBP such as toxins,

CC immunomodulating agents

CC The immunoglobulin c

CC facilitates recovery

cell surface glycoprotein

XX 270 111

Query Match

Matches 34; Conserva

QY 1 CQPWSCSGHGECEI

Db 158 cqpwscsyhgcevei


```

RESULT 3
AAR98128
ID AAR98128 standard; Protein: 370 AA.
XX
AC AAR98128;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Region 1..37
FH FT /label= Signal region.
FH FT 38..154
FH FT /label= Lectin domain.
FH FT 159..192
FH FT /label= EGF domain.
FH FT 196..256
FH FT /label= Complement binding repeat 1.
FH FT 257..315
FH FT /label= Complement binding repeat 2.
FH FT 331..353
FH FT /label= Transmembrane domain.
FH FT 354..370
FH FT /label= Cytoplasmic domain.
XX
PN US5514582-A.
XX
PD 07-MAY-1996.
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
XX immunoglobulin constant region. The receptor is not a member of the
XX immunoglobulin super family, nor a multiple subunit polypeptide
XX encoded by discrete genes. The hybrid immunoglobulin chain combines
XX the adhesion/targeting of a ligand binding partner (LBP) with the
XX effector functions of immunoglobulin and can bind to and/or activate
XX more than one ligand. It can be used diagnostically for the in
XX vitro assay of LBP and their targets; or therapeutically to deliver
XX LBP such as toxins, enzymes, growth factors to particular cells.
XX Typical applications are as antiviral, neuromodulating and
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX The immunoglobulin component increases plasma half life and
XX facilitates purification while deletion of the transmembrane region

```

```

CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant is a Ser220 deletion mutant.
XX
SQ Sequence 370 AA;
XX
Query Match 100.0%; Score 215; DB 17; Length 370;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 COPWSCSGHGECEIINNHNTCNCDYGYGPOCQL 34
DB 159 cqpwscsghgeceiinhntcncdvygygpcqql 192
XX
RESULT 4
AAR98130
ID AAR98130 standard; Protein: 370 AA.
XX
AC AAR98130;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Region 1..37
FH FT /label= Signal region.
FH FT 38..154
FH FT /label= Lectin domain.
FH FT 159..192
FH FT /label= EGF domain.
FH FT 196..257
FH FT /label= Complement binding repeat 1.
FH FT 258..315
FH FT /label= Complement binding repeat 2.
FH FT 331..353
FH FT /label= Transmembrane domain.
FH FT 354..370
FH FT /label= Cytoplasmic domain.
XX
PN US5514582-A.
XX
PD 07-MAY-1996.
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.

```

XX A hybrid immunoglobulin chain comprising the ligand binding site of
 CC a single transmembrane receptor without an active transmembrane
 CC region; fused at its C-terminus with the N-terminus of an
 CC immunoglobulin constant region. The receptor is not a member of the
 CC immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC the adhesion/targeting of a ligand binding partner (LBP) with the
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver
 CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery, improves aqueous solubility and removes
 CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-R98135. This variant is an Ile296 deletion mutant.

XX Sequence 370 AA:

Query Match 100.0%; Score 215; DB 17; Length 370;

Best Local Similarity 100.0%; Pred. No. 3e-14; Mismatches 0; Indels 0; Gaps 0;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CQPMSCSGHGECEVEIINHTCNCNDVGYGPQCL 34
 Db 159 cqpmscsgheceveiinhtcncndvgygpqcl 192

RESULT 5

AAR98109 standard; Protein; 371 AA.

AC AAR98109;

DT 31-OCT-1996 (first entry)

DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

KW Immunoglobulin: transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

KW metastasis.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Region 1..37 /label= Signal region.

FT Domain 38..154 /label= Lectin domain.

FT Domain 159..192 /label= EGF domain.

FT Binding-site 196..257 /label= Complement binding repeat 1.

FT Binding-site 258..316 /label= Complement binding repeat 2.

FT Domain 332..354 /label= Transmembrane domain.

FT Domain 355..371 /label= Cytoplasmic domain.

PN US5514582-A.

PD 07-MAY-1996.

PF 23-FEB-1989; 89US-0315015.

PR 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.

PA (GENTH) GENENTECH INC.

PI Capon DJ, Lasky LA;

DR WPI, 1996-238773/24.

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand

PT binding site of a receptor fused to Ig constant region - useful for

PT diagnosis and treatment e.g. of inflammation

PS Disclosure; Page 19; 41pp; English.

CC A hybrid immunoglobulin chain comprising the ligand binding site of

CC a single transmembrane receptor without an active transmembrane

CC region; fused at its C-terminus with the N-terminus of an

CC immunoglobulin constant region. The receptor is not a member of the

CC immunoglobulin super family, nor a multiple subunit polypeptide

CC encoded by discrete genes. The hybrid immunoglobulin chain combines

CC the adhesion/targeting of a ligand binding partner (LBP) with the

CC effector functions of immunoglobulin and can bind to and/or activate

CC more than one ligand. It can be used diagnostically for the in

CC vitro assay of LBP and their targets; or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.

CC Typical applications are as antiviral, neuromodulating and

CC immunomodulating agents, or as modulators of cell adhesion (e.g. in

CC treating graft rejection; inflammation; metastasis of lymphoma etc.)

CC The immunoglobulin component increases plasma half life and

CC facilitates purification while deletion of the transmembrane region

CC facilitates recovery, improves aqueous solubility and removes

CC potentially immunogenic epitopes. Variants of the human lymphocyte

CC cell surface glycoprotein described in AAR98106 are given in

CC AAR98109-R98135. This variant contains a Arg58-Asp59; Lys-Glu

CC substitution.

SO Sequence 371 AA;

Query Match 100.0%; Score 215; DB 17; Length 371;

Best Local Similarity 100.0%; Pred. No. 3e-14; Mismatches 0; Indels 0; Gaps 0;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQPMSCSGHGECEVEIINHTCNCNDVGYGPQCL 34

Db 159 cqpmscsgheceveiinhtcncndvgygpqcl 192

RESULT 6

AAR98110 standard; Protein; 371 AA.

AC AAR98110;

DT 01-NOV-1996 (first entry)

DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

KW Immunoglobulin: transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

KW metastasis.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Region 1..37 /label= Signal region.

FT Domain 38..154

FT		/label- Lectin domain.
FT	Domain	159..192
FT		/label- EGF domain.
FT	Binding-site	196..257
FT		/label- Complement binding repeat 1.
FT	Binding-site	258..316
FT		/label- Complement binding repeat 2.
FT	Domain	332..354
FT		/label- Transmembrane domain.
FT	Domain	355..371
FT		/label- Cytoplasmic domain.
PX		
PN	USS514582-A.	
PX		
PD	07-MAY-1996.	
PX		
PF	23-FEB-1989;	89US-0315015.
PX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
PX		
PA	(GETH) GENENTECH INC.	
PI	Capon DJ, Lasky LA:	
DR	WPI: 1996-238773/24.	
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation	
PS	Disclosure: Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region; fused at its C-terminus with the N-terminus of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targeting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the in vitro assay of LBP and their targets, or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particular cells. Typical applications are as antiviral, neuromodulating and immunomodulating agents, or as modulators of cell adhesion (e.g. in treating graft rejection; inflammation; metastasis of lymphoma etc.) The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR98106 are given in AAR98109-R98135. This variant contains an Ala71Ser substitution.	
SQ	Sequence 371 AA:	
	Query Match 100.0%; Score 215; DB 17; Length 371;	
	Best Local Similarity 100.0%; Pident. No. 3e-14;	
	Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 COPWSCSGHGECVELIINNHTCNCNDVGYYGPCOL 34	
DB	159 cqpwscsghgecvelinnhtcncndvgyygpqcl 192	
RESULT 7		
AAR98111		
ID	AAR98111 standard; Protein: 371 AA.	
XX		

AC	AAK98111;	
XX		
DT	01-NOV-1996	(first entry)
XX		
DE	Human lymphocyte cell surface glycoprotein (HULHR) variant.	
XX		
KM	Immunoglobulin; transmembrane receptor; adhesion; targeting;	
KM	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KM	immunomodulator; cell adhesion; graft rejection; inflammation;	
KM	metastasis.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= EGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1.
FT	Binding-site	258..316
FT		/label= Complement binding repeat 2.
FT	Domain	332..354
FT		/label= Transmembrane domain.
FT	Domain	355..371
FT		/label= Cytoplasmic domain.
XX		
PN	US5514582-A.	
XX		
PD	07-MAY-1996.	
XX		
XX	23-FEB-1989;	89US-0315015.
PF		
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GENTH) GENENTECH INC.	
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI, 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
XX		
PS	Disclosure: Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region, fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAK98106 are given in	
CC	AAK98109-988135. This variant contains an Lys78Gln substitution.	

```
XX Sequence 371 AA;
SQ
Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 COPWSCSGHGECEVEIINNHTCNCDCVGYGPOCOL 34
Db 159 cqpwsesghgeceveinnhtcncdcvgygpcql 192
|||||
RESULT 8
AAR98112
ID AAR98112 standard; Protein; 371 AA.
XX
AC AAR98112;
XX
XX 01-NOV-1996 (first entry)
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX immunomodulator; cell adhesion; graft rejection; inflammation;
XX metastasis.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..37
XX FT /label= Signal region.
XX FT 38..154
XX FT /label= Lectin domain.
XX FT 159..192
XX FT /label= EGF domain.
XX FT 196..257
XX FT /label= Complement binding repeat 1.
XX FT 258..316
XX FT /label= Complement binding repeat 2.
XX FT 332..354
XX FT /label= Transmembrane domain.
XX FT 355..371
XX FT /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
XX PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to its constant region - useful for
XX diagnosis and treatment e.g. of inflammation
XX
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
```

```
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targetting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an Asp116Glu substitution.
CC
XX
XX Sequence 371 AA;
SQ
Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 COPWSCSGHGECEVEIINNHTCNCDCVGYGPOCOL 34
Db 159 cqpwsesghgeceveinnhtcncdcvgygpcql 192
|||||
RESULT 9
AAR98113
ID AAR98113 standard; Protein; 371 AA.
XX
AC AAR98113;
XX
XX 01-NOV-1996 (first entry)
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX immunomodulator; cell adhesion; graft rejection; inflammation;
XX metastasis.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..37
XX FT /label= Signal region.
XX FT 38..154
XX FT /label= Lectin domain.
XX FT 159..192
XX FT /label= EGF domain.
XX FT 196..257
XX FT /label= Complement binding repeat 1.
XX FT 258..316
XX FT /label= Complement binding repeat 2.
XX FT 332..354
XX FT /label= Transmembrane domain.
XX FT 355..371
XX FT /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
```

```

PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
PA
XX Capon DJ, Lasky LA;
PI
XX WPI: 1996-238773/24.
DR
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
PS
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Leu150Val substitution.
XX
XX Sequence 371 AA;
SQ
Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 COPWSCSGHCEVEIINNHTCNCDDVGYGPOCOL 34
DB 159 CQPWSCSGHCEVEIINNHTCNCDDVGYGPOCQL 192
RESULT 10
AAR98117
ID AAR98117 standard; Protein: 371 AA.
XX
XX AAR98117;
AC
XX 01-NOV-1996 (first entry)
DT
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
DE
XX
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT Region 1..37
FT Domain /label= Signal region.
FT Domain 38..154
FT Domain /label= Lectin domain.
FT Domain 159..192
FT Domain /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.

```

```

FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT Domain /label= Transmembrane domain.
FT Domain 355..371
FT Domain /label= Cytoplasmic domain.
PN US5514582-A.
XX
XX 07-MAY-1996.
PD
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
PA
XX Capon DJ, Lasky LA;
PI
XX WPI: 1996-238773/24.
DR
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
PS
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Thr151Ser substitution.
XX
XX Sequence 371 AA;
SQ
Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 COPWSCSGHCEVEIINNHTCNCDDVGYGPOCOL 34
DB 159 CQPWSCSGHCEVEIINNHTCNCDDVGYGPOCQL 192
RESULT 11
AAR98118
ID AAR98118 standard; Protein: 371 AA.
XX
XX AAR98118;
AC
XX 01-NOV-1996 (first entry)
DT
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
DE

```

```

XX      Immunoglobulin; transmembrane receptor; adhesion; targetting;
KM      diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KM      immunomodulator; cell adhesion; graft rejection; inflammation;
KM      metastasis;
XX      Homo sapiens.
OS
XX      Key
FH      Location/Qualifiers
FT      1..37
FT      /label= Signal region.
FT      Domain
FT      38..154
FT      /label= Lectin domain.
FT      Domain
FT      159..192
FT      /label= EGF domain.
FT      Binding-site
FT      196..257
FT      /label= Complement binding repeat 1.
FT      Binding-site
FT      258..316
FT      /label= Complement binding repeat 2.
FT      Domain
FT      332..354
FT      /label= Transmembrane domain.
FT      Domain
FT      355..371
FT      /label= Cytoplasmic domain.
XX
XX      US5514582-A.
XX
XX      07-MAY-1996.
XX
XX      23-FEB-1989; 89US-0315015.
XX
XX      22-NOV-1989; 89US-0440625.
XX      23-FEB-1989; 89US-0315015.
XX      16-DEC-1991; 91US-0808122.
XX      08-DEC-1992; 92US-0986931.
XX      21-JAN-1994; 94US-0185670.
XX
XX      (GETH ) GENENTECH INC.
XX
XX      Capon DJ, Lasky LA;
XX
XX      WPI; 1996-238773/24.
XX
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT      binding site of a receptor fused to Ig constant region - useful for
PT      diagnosis and treatment e.g. of inflammation
XX
XX      Disclosure; Page 19; 41pp; English.
XX
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
CC      a single transmembrane receptor without an active transmembrane
CC      region; fused at its C-terminus with the N-terminus of an
CC      immunoglobulin constant region. The receptor is not a member of the
CC      immunoglobulin super family, nor a multiple subunit polypeptide
CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
CC      the adhesion/targetting of a ligand binding partner (LBP) with the
CC      effector functions of immunoglobulin and can bind to and/or activate
CC      more than one ligand. It can be used diagnostically for the in
CC      vitro assay of LBP and their targets; or therapeutically to deliver
CC      LBP such as toxins, enzymes, growth factors to particular cells.
CC      Typical applications are as antiviral, neuromodulating and
CC      immunomodulating agents; or as modulators of cell adhesion (e.g. in
CC      treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC      The immunoglobulin component increases plasma half life and
CC      facilitates purification while deletion of the transmembrane region
CC      facilitates recovery. Improves aqueous solubility and removes
CC      potentially immunogenic epitopes. Variants of the human lymphocyte
CC      cell surface glycoprotein described in AAR98106 are given in
CC      AAR98109-938135. This variant contains a Phe214Leu substitution.
XX
XX      Sequence 371 AA;
XX
XX      Query Match 100.0%; Score 215; DB 17; Length 371;

```

```

Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      1 CQPMSCSGHGECYEITNNHTCNDVGYYGPQCL 34
Db      159 cqpwsesghgecvellnhhtcndvgyypqcl 192

RESULT 12
AAR98119
ID      AAR98119 standard; Protein; 371 AA.
XX
XX      AAR98119;
AC
XX
XX      01-NOV-1996 (first entry)
DT
XX
XX      Human lymphocyte cell surface glycoprotein (HuLHR) variant.
DE
XX
XX      Immunoglobulin; transmembrane receptor; adhesion; targetting;
KM      diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KM      immunomodulator; cell adhesion; graft rejection; inflammation;
KM      metastasis.
XX
XX      Homo sapiens.
XX
XX      Key
FH      Location/Qualifiers
FT      1..37
FT      /label= Signal region.
FT      Domain
FT      38..154
FT      /label= Lectin domain.
FT      Domain
FT      159..192
FT      /label= EGF domain.
FT      Binding-site
FT      196..257
FT      /label= Complement binding repeat 1.
FT      Binding-site
FT      258..316
FT      /label= Complement binding repeat 2.
FT      Domain
FT      332..354
FT      /label= Transmembrane domain.
FT      Domain
FT      355..371
FT      /label= Cytoplasmic domain.
XX
XX      US5514582-A.
XX
XX      07-MAY-1996.
XX
XX      23-FEB-1989; 89US-0315015.
XX
XX      22-NOV-1989; 89US-0440625.
XX      23-FEB-1989; 89US-0315015.
XX      16-DEC-1991; 91US-0808122.
XX      08-DEC-1992; 92US-0986931.
XX      21-JAN-1994; 94US-0185670.
XX
XX      (GETH ) GENENTECH INC.
XX
XX      Capon DJ, Lasky LA;
XX
XX      WPI; 1996-238773/24.
XX
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT      binding site of a receptor fused to Ig constant region - useful for
PT      diagnosis and treatment e.g. of inflammation
XX
XX      Disclosure; Page 19; 41pp; English.
XX
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
CC      a single transmembrane receptor without an active transmembrane
CC      region; fused at its C-terminus with the N-terminus of an
CC      immunoglobulin constant region. The receptor is not a member of the
CC      immunoglobulin super family, nor a multiple subunit polypeptide
CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
CC      the adhesion/targetting of a ligand binding partner (LBP) with the
CC      effector functions of immunoglobulin and can bind to and/or activate

```

CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Ser226Thr substitution.
CC
XX
SO Sequence 371 AA;

Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGEVETIINHTCNCDDVGYGPGCQL 34
Db 159 CQPMSCSGHGEVETIINHTCNCDDVGYGPGCQL 192

RESULT 13
AAR98120
ID AAR98120 standard; Protein; 371 AA.
XX
AC AAR98120;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
KM
XX Homo sapiens.
XX OS
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371
FT /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX PD 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 16-DEC-1991; 91US-0808122.
XX 08-DEC-1992; 92US-0986931.
XX 21-JAN-1994; 94US-0185670.
XX
XX (GETH) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;

XX
DR WPI: 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure: Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Phe244Met substitution.
CC
XX
SO Sequence 371 AA;

Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGEVETIINHTCNCDDVGYGPGCQL 34
Db 159 CQPMSCSGHGEVETIINHTCNCDDVGYGPGCQL 192

RESULT 14
AAR98121
ID AAR98121 standard; Protein; 371 AA.
XX
AC AAR98121;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
KM
XX Homo sapiens.
XX OS
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371

	/Label= Cytoplasmic domain.
FT	
FN	US5514582-A.
PD	07-MAY-1996.
XX	
PF	23-FEB-1989; 89US-0315015.
PR	22-NOV-1989; 89US-0440625.
PR	23-FEB-1989; 89US-0315015.
PR	16-DEC-1991; 91US-0808122.
PR	08-DEC-1992; 92US-0986931.
PR	21-JAN-1994; 94US-0185670.
XX	
PA	(GETH) GENENTECH INC.
XI	
PI	Capon DJ, Lasky LA;
DR	WPI; 1996-238773/24.
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation
PS	Disclosure; Page 19; 41pp; English.
CC	A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region; fused at its C-terminals with the N-terminals of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targetting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the in vitro assay of LBP and their targets; or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particular cells. Typical applications are as antiviral, neuromodulating and immunomodulating agents, or as modulators of cell adhesion (e.g., In treating graft rejection; inflammation; metastasis of lymphoma etc.) The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR98106 are given in AAR98109-R98135. This variant contains a Thr282Ser substitution.
SQ	Sequence 371 AA;
Query Match	100.0%; Score 215; DB 17; Length 371;
Best Local Similarity	100.0%; Pval. No. 3e-14;
Matches	34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y	1 COPWSCSGHGCEVEETINNHTGCNCDDVGYYGPCOL_34
Df	159 cgpwscshgceveeinhtcndvgyypqcl_192
RESULT	15
AAR98122	ID ID AAR98122 standard; Protein; 371 AA.
AC	AAR98122;
DT	01-NOV-1996 (first entry)
DE	Human lymphocyte cell surface glycoprotein (HLHR) variant.
KM	Immunoglobulin; transmembrane receptor; adhesion; targetting; diagnosis; therapy; drug delivery; antiviral; neuromodulator; immunostimulatory; cell adhesion; graft rejection; inflammation; metastatis.

XX	Homo sapiens.
OS	
XX	
XX	
FH	Key
FT	Region
FT	Domain
FT	Domain
FT	Binding-site
FT	Binding-site
FT	Binding-site
FT	Domain
FT	Domain
FT	Domain
FN	US5514582-A.
PD	07-MAY-1996.
XX	
PP	23-FEB-1989;
XX	
PR	22-NOV-1989;
PR	23-FEB-1989;
PR	16-DEC-1991;
PR	08-DEC-1992;
PR	21-JAN-1994;
XX	
PA	(GETH) GENENTECH INC.
PI	Capon DJ, Lasky LA:
DR	WPI; 1996-238773/24.
XX	
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation
PT	
PS	Disclosure; Page 19; 41pp; English.
CC	A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region; fused at its C-terminus with the N-terminus of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targeting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the in vitro assay of LBP and their targets; or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particular cells. Typical applications are as antiviral, neuromodulating and immunomodulating agents, or as modulators of cell adhesion (e.g. in treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC	The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR98106 are given in AAR98109-R98135. This variant contains an Ile28Val substitution.
SO	Sequence 371 AA;
Query Match	100.0%; Score 215; DB 17; Length 371;
Best Local Similarity	100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
1 COPMSCGHECVETLNHTNCMDYGYRPOCOL 34 	

Db 159 cqpwscsghecvellmhtcndvgyygpqcq1 192

RESULT 16

AAR98123

ID AAR98123 standard; Protein: 371 AA.

XX

AC AAR98123:

XX

DT 01-NOV-1996 (first entry)

XX

DE Human lymphocyte cell surface glycoprotein (HLHR) variant.

XX

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

KW metastasis.

XX

OS Homo sapiens.

XX

PH Key

FT Region

FT Domain

FT Domain

FT Domain

FT Binding-site

FT Binding-site

FT Binding-site

FT Binding-site

FT Domain

FT Domain

FT Domain

PN US5514582-A.

PD 07-MAY-1996.

XX

PF 23-FEB-1989; 89US-0315015.

XX

PR 22-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

XX

PA (GETH) GENENTECH INC.

XX

PI Capon DJ, Lasky LA;

XX

DR WPI; 1996-238773/24.

XX

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand

PT binding site of a receptor fused to Ig constant region - useful for

PT diagnosis and treatment e.g. of inflammation

XX

PS Disclosure; Page 19; 41pp; English.

XX

XX

CC A hybrid immunoglobulin chain comprising the ligand binding site of

CC a single transmembrane receptor without an active transmembrane

CC region; fused at its C-terminus with the N-terminus of an

CC immunoglobulin constant region. The receptor is not a member of the

CC immunoglobulin super family, nor a multiple subunit polypeptide

CC encoded by discrete genes. The hybrid immunoglobulin chain combines

CC the adhesion/targeting of a ligand binding partner (LBP) with the

CC effector functions of immunoglobulin and can bind to and/or activate

CC more than one ligand. It can be used diagnostically for the in

CC vitro assay of LBP and their targets; or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.

CC Typical applications are as antiviral, neuromodulating and

CC immunomodulating agents, or as modulators of cell adhesion (e.g. in

CC treating graft rejection; inflammation; metastasis of lymphoma etc.)

CC The immunoglobulin component increases plasma half life and

CC facilitates purification while deletion of the transmembrane region

CC facilitates recovery, improves aqueous solubility and removes

CC potentially immunogenic epitopes. Variants of the human lymphocyte

CC cell surface glycoprotein described in AAR98106 are given in

CC AAR98109-R98135. This variant contains a Lys298-Lys299; Arg-Arg

CC substitution.

XX

XX Sequence 371 AA;

SQ

Query Match 100.0%; Score 215; DB 17; Length 371;

Best Local Similarity 100.0%; Pred. No. 3e-14;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPWSCSGHECVELLMHTCNDVGYYPQCQL 34

Db 159 cqpwscsghecvellmhtcndvgyygpqcq1 192

|||||

RESULT 17

AAR98124

ID AAR98124 standard; Protein: 371 AA.

XX

AC AAR98124:

XX

DT 01-NOV-1996 (first entry)

XX

DE Human lymphocyte cell surface glycoprotein (HLHR) variant.

XX

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

KW metastasis.

XX

OS Homo sapiens.

XX

PH Key

FT Region

FT Domain

FT Domain

FT Domain

FT Binding-site

FT Binding-site

FT Binding-site

FT Binding-site

FT Domain

FT Domain

FT Domain

PN US5514582-A.

PD 07-MAY-1996.

XX

PF 23-FEB-1989; 89US-0315015.

XX

PR 22-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

XX

PA (GETH) GENENTECH INC.

XX

PI Capon DJ, Lasky LA;

XX

DR WPI; 1996-238773/24.

XX

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand

binding site of a receptor fused to Ig constant region - useful for
diagnosis and treatment e.g. of inflammation

Disclosure: Page 19; 41pp; English.

A hybrid immunoglobulin chain comprising the ligand binding site of
a single transmembrane receptor without an active transmembrane
region; fused at its C-terminus with the N-terminus of an
immunoglobulin constant region. The receptor is not a member of the
immunoglobulin super family, nor a multiple subunit polypeptide
encoded by discrete genes. The hybrid immunoglobulin chain combines
the adhesion/targeting of a ligand binding partner (LBP) with the
effector functions of immunoglobulin and can bind to and/or activate
more than one ligand. It can be used diagnostically for the *in*
vitro assay of LBP and their targets; or therapeutically to deliver
LBP such as toxins, enzymes, growth factors to particular cells.
Typical applications are as antiviral, neuromodulating and
immunomodulating agents, or as modulators of cell adhesion (e.g. in
treating graft rejection; inflammation; metastasis of lymphoma etc.)
The immunoglobulin component increases plasma half life and
facilitates purification while deletion of the transmembrane region
facilitates recovery, improves aqueous solubility and removes
potentially immunogenic epitopes. Variants of the human lymphocyte
cell surface glycoprotein described in AAP8106 are given in
AAP898109-R98135. This variant contains an Ile302Leu substitution.

SQ Sequence 371 AA;

Query Match	100.0%	Score 215; DB 17	Length 371;
Best Local Similarity	100.0%	Pred. No. 3e-14;	
Matches 34; Conservative	0;	Mismatches	0; Gaps 0;

QY 1 CQPMSCSGHGECEVEIINNHTCNCNDGYRPGCQL 34
|||||
Db 159 cqpmscsghgeceveIinnhtcncndgyrpgcql 192

RESULT	18
AAK98129	
ID	AAK98129 standard; Protein; 371 AA.
XX	
AC	AAK98129;

DT 01-NOV-1996 (first entry)

Human lymphocyte cell surface glycoprotein (HULHR) variant

KM Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.

OS Homo sapiens.

FT	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= EGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1.
FT	Binding-site	258..315
FT		/label= Complement binding repeat 2.
FT	Domain	331..353
FT		/label= Transmembrane domain.
FT	Domain	354..370
FT		/label= Cytoplasmic domain.

PN US5514582-A.
XX

PD	07-MAY-1996.
XX	
PF	23-FEB-1989; 89US-0315015.

PR 22-NOV-1989; 89US-0440625;
PR 23-FEB-1989; 89US-0315015;
PR 16-DEC-1991; 91US-0808122;
PR 08-DEC-1992; 92US-0986531;
PR 21-JAN-1994; 94US-0185670.
PR

PA (GETH) GENENTECH INC.

PI Capon DJ, Lasky LA,
XX
DR WPI; 1996-238773/24.

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation

PS Disclosure; Page 19; 41pp; English

CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region, fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targetting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the *in*
CC *vitro* assay of LBP and their targets, or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AA098106 are given in
CC AA098109-098135. This variant is a Asn271 deletion mutant.

Sequence 371 AA;

Query Match	100.0%;	Score 215;	DB 17;	Length 371;
Best Local Similarity	100.0%;	Pred. No. 3e-14;		
Matches	34;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

QY 1 CQPMSCSGHGECVELINHTCMCDVGYGPOCQL 34
|||
Db 160 CQPMSCSGHGECEVLINHTCMCDVGYGPQCQL 193

RESULT 19
AAR12469
ID AAR12469 standard; Protein; 372 AA

AC AAR12469;

DT 09-SEP-1991 (first entry)

Human lymphocyte cell surface glycoprotein (LHR).

KW Ligand binding partner; lbp; stable plasma protein; spp; antiviral;
 KW immunomodulatory; receptor mediated abnormality.

OS Homo sapiens.

Key	Location/Qualifiers
FH	20..32
FT	/label= signal peptide
FT	


```

XX Key Location/Qualifiers
FH Peptide 1..19
FT Peptide 20..38
FT Peptide /label= signal
FT Modified-site 60..62
FT Modified-site /label= potential N-linked glycosylation site
FT Modified-site 104..106
FT Modified-site /label= see above
FT Modified-site 177..179
FT Modified-site /label= see above
FT Modified-site 216..218
FT Modified-site /label= see above
FT Modified-site 232..234
FT Modified-site /label= see above
FT Modified-site 271..273
FT Modified-site /label= see above
FT Modified-site 311..313
FT Modified-site /label= see above
FT Region 333..355
FT Region /label= stop transfer sequence
XX US5116964-A.
XX 26-MAY-1992.
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 22-NOV-1989; 89US-0440625.
XX (GETH ) GENENTECH INC.
XX Capon DJ, Lasky LA;
XX WPI: 1992-199589/24.
XX N-PSDB; AAQ24987.
XX Nucleic acid encoding polypeptide fusions - comprising ligand
XX binding partner protein and immunoglobulin chain, for use in
XX diagnosis and therapy
XX Disclosure; Fig 1-1 - 1-3; 43pp; English.
XX LHR mediates the binding of lymphocytes to the endothelium of
XX lymphoid tissue. Full length cDNA clones and DNA encoding the human
XX and the murine LHR (HuLHR and MLHR, respectively) have been
XX identified and isolated (see AAQ24987 and AAQ24988). LHR is a
XX glycoprotein which contains the following protein domains: a signal
XX sequence, a carbohydrate binding domain, and epidermal growth
XX factor-like (egf) domain, at least one and preferably two complement
XX binding domain repeat, a transmembrane binding domain (TMD), and a
XX charged intracellular or cytoplasmic domain. LHR is used as the
XX ligand-binding partner in fusion polypeptides with an immunoglobulin,
XX for use in diagnosis and therapy.
XX
XX Sequence 372 AA:
SO

```

```

Query Match 100.0%; Score 215; DB 13; Length 372;
Best local similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 CQPMSCSGHGEVVEIINHTCNCVGYGPOCOL 34
    ||||||||||||||||||||||||||||
Db 160 cqpmscsghgevcveiinhtcncdvgygpcql 193

```

```

RESULT 22
AAK37960
ID AAK37960 standard; Protein; 372 AA.
XX
AC AAK37960;

```

```

XX 08-OCT-1993 (first entry)
DT Human Lymphocyte Homing Receptor.
XX HuLHR; Lymphocyte binding inhibition; lymphoma metastasis;
XX transplant rejection; inflammation.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..38
FT Protein /label= signal_sequence
FT Domain /note= "Trp39 is probable N-terminus of mature LHR"
FT Modified-site 39..155
FT Modified-site /label= Lectin_domain
FT Modified-site 60..62
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 104..106
FT Modified-site /note= "potential N-glycosylation site"
FT Domain 160..193
FT Domain /label= EGF_domain
FT Modified-site 177..179
FT Region /note= "potential N-glycosylation site"
FT Region 197..258
FT Modified-site /label= Complement_Binding_Repeat_1
FT Modified-site 216..218
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 232..234
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 246..248
FT Modified-site /note= "potential N-glycosylation site"
FT Region 259..317
FT Modified-site /label= Complement_Binding_Repeat_2
FT Modified-site 271..273
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 311..313
FT Modified-site /note= "potential N-glycosylation site"
FT Domain 333..355
FT Domain /label= Transmembrane_Domain
FT Domain /note= "stop transfer sequence"
FT Domain 356..372
FT Domain /label= Cytoplasmic_Domain
XX US5216131-A.
XX 01-JUN-1993.
XX 23-FEB-1989; 89US-0315015.
XX 23-FEB-1989; 89US-0315015.
XX 23-FEB-1989; 89US-0315015.
XX 31-OCT-1991; 91US-0786149.
XX (GETH ) GENENTECH INC.
XX
XX Lasky LA, Rosen SD, Singer MS, Stachel SE, Yednock TA;
XX WPI: 1993-188588/23.
XX N-PSDB; AAQ43154.
XX Human and murine lymphocyte homing receptors to treat graft
XX rejection and inflammation - comprise carbohydrate binding,
XX epidermal growth factor and complement binding domains
XX Claim 1; Fig 1 and Fig 3; 32pp; English.
XX
XX A human peripheral blood lymphocyte cDNA library in lambda gt10 was
XX screened with a 2.2kb EcoRI insert of the murine Mcl14 antigen clone
XX (i.e. a murine LHR sequence). The largest EcoRI insert (2.2kb) was
XX isolated and sequenced. The ORF codes for 372 amino acids with a mol.
XX wt. of approximately 42,200. Comparison of the HuLHR amino acid
XX sequence with the murine LHR sequence (AAK37961) showed a high degree

```

CC of amino acid conservation in each of the LHR domains, e.g. 96% in
 CC the transmembrane domain and 83% in the carbohydrate binding domain.
 CC The LHRs could be used to compete with the normal binding of
 CC lymphocytes to lymphoid tissue to treat inflammation or graft
 CC rejection. They could also be used to control lymphoma metastasis
 CC and to treat conditions involving lymphocyte accumulation.

XX Sequence 372 AA:

Query Match 100.0%; Score 215; DB 14; Length 372;
 Best Local Similarity 100.0%; Pred. No. 3e-14;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CQPMSCSGHGECEVEIINHTCNCDDVGYGPQCQL 34
 Db 160 cqpmscsghgeceveilmhtcncddvgygpqcql 193

RESULT 23

ID AAR38908 standard; Protein; 372 AA.

XX AAR38908;

DT 11-JAN-1994 (first entry)

DE HuLHR.

XX Human; murine; lymphocyte; cell surface glycoprotein; homing receptor;
 KW LHR; endothelium; lymphoid tissue; signal; domain; complement binding;
 KW carbohydrate binding; epidermal growth factor-like; egf; intracellular;
 KW transmembrane binding; cytoplasmic; ligand binding partner protein;
 KW TM; LBP.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 20..32 /note= "Signal peptide"

FT Protein 39..372 /note= "Mature protein"

FT Domain 39..155 /note= "Lectin domain"

FT Domain 160..193 /note= "egf domain"

FT Domain 197..317 /note= "Complement factor binding domain"

FT Domain 333..355 /note= "Transmembrane binding domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

FT Domain 356..372 /note= "Cytoplasmic domain"

XX Disclosure: Fig 1: 44pp: English.

XX The sequences given in AAR38908-09 represent human and murine lymphocyte
 CC cell surface glycoprotein (LHR) respectively. These proteins mediate
 CC the binding of lymphocytes to the endothelium of lymphoid tissue. LHR
 CC is a glycoprotein which contains a signal domain, a carbohydrate
 CC binding domain, an epidermal growth factor-like (egf) domain, at least
 CC one complement binding domain repeat, a transmembrane binding domain
 CC (TM) and a charged intracellular or cytoplasmic domain. The murine
 CC and human amino acid sequences show a high degree of overall homology
 CC (83%), however degrees of homology between the various domains is
 CC variable. These proteins may be fused to a ligand binding partner
 CC protein (LBP) which causes an increase in the half life of the LHR.
 CC The fusions may be used therapeutically to compete with the normal
 CC binding of lymphocytes to lymphoid tissue. They may be used in organ
 CC or graft rejection and for the treatment of inflammation.

XX Sequence 372 AA:

Query Match 100.0%; Score 215; DB 14; Length 372;
 Best Local Similarity 100.0%; Pred. No. 3e-14;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CQPMSCSGHGECEVEIINHTCNCDDVGYGPQCQL 34
 Db 160 cqpmscsghgeceveilmhtcncddvgygpqcql 193

RESULT 24

ID AAR76506 standard; Protein; 372 AA.

XX AAR76506;

DT 01-DEC-1995 (first entry)

DE Human LHR.

XX Lymphocyte homing receptor; lymphocyte cell surface glycoprotein;
 KW LHR; ligand binding partner; immunoglobulin; constant region;
 KW antibody engineering; immunomodulator.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

FT Peptide 20..32 /label= "Sig.peptide"

```

FT      /note= "putative stop transfer or membrane anchor
FT      domain"
FT      Region          358..372
FT      FT              /note= "putative intracellular region"
XX      PN              US5428130-A.
XX      PD              27-JUN-1995.
XX      PF              23-FEB-1989;    89US-0315015.
XX      PR              22-NOV-1989;    89US-0440625.
XX      PR              23-FEB-1989;    89US-0315015.
PR      16-DEC-1991;    91US-0808122.
XX      PR              08-DEC-1992;    92US-0986931.
XX      PA              (GERTH ) GENENTECH INC.
XX      PI              Capon DJ, Lasky LA;
XX      DR              WPI, 1995-240086/31.
XX      DR              N-PsDB; AAQ92802.
XX      PT              New hybrid ligand binding partner molecules - fused to immunoglobulin
PT      constant region sequences to increase stability and in vivo plasma
PS      half-life
PS      Disclosure; Fig.1a-1c; 40pp; English.
XX      CC              A murine Mel 14 antigen cDNA clone was used to screen a lambda gt10
CC      cDNA library derived from human peripheral blood lymphocyte mRNA
XX      obdI. from primary cells. A cDNA clone encoding LHR was isolated.
SQ      Sequence       372 AA:

Query Match           100.0%; Score 215; DB 16; Length 372;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches   34; Conservative     0; Mismatches     0; Indels     0; Gaps     0;

OY      1 CQPWSCSGHGECVELINNHTCNCDDVGYYGPQCQL 34
        |||
D6      160 CQPWSCSGHGECVELINNHtcncddvgyypqcql 193

RESULT  25
ID      AAR83050 standard; Protein; 372 AA.
XX      AC      AAR83050;
DT      DT      31-JAN-1996 (first entry)
XX      DE      Human LHR.
XX      KW      Lymphocyte cell surface glycoprotein; LHR; transmembrane receptor;
KW      immunoglobulin; IgG; constant region; receptor-mediated disease;
XX      KM      vector; plasma-1lfe.
XX      OS      Homo sapiens.
XX      FH      Key Location/Qualifiers
FH      FT      Peptide 20..32
FT      FT      /label= Sig-peptide
FT      FT      Domain 39..155
FT      FT      /label= Carbohydrate_binding_domain
FT      FT      Domain 160..193
FT      FT      /label= Epidermal_growth_factor_domain
FT      FT      Domain 197..317
FT      FT      /label= Complement_factor_binding_domain
FT      FT      Domain 333..355
FT      FT      /label= Transmembrane_binding_domain
FT      FT      Domain 356..372

```

FT		/label= Cytoplasmic_domain
FT	Modified-site	/note="potential stop transfer sequence"
FT	Modified-site	60..62
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	104..106
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	177..179
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	216..218
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	232..234
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	248..248
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	271..273
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	311..313
FT	Modified-site	/label= N-glycosylation_site
XX		
PN	US5455165-A.	
XX		
XX	03-OCT-1995.	
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185669.
XX		
PA	(GETH) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI; 1995-350776/45.	
XX	N-PSDB; AAT05869.	
PT	Expression vector encoding fusion protein to increase plasma life -	
PT	comprises receptor ligand binding site and Ig constant region, for	
PT	treatment of receptor mediated disease	
XX		
XX	Disclosure; Fig 1; 42pp: English.	
XX		
CC	A mouse LHR (lymphocyte cell surface glycoprotein) cDNA clone was used	
CC	to screen an oligo-dT primed lambda gt10 cDNA library derived from	
CC	human peripheral blood lymphocyte mRNA obtd. from primary cells. A	
CC	2.2 kb clone (sequence given in AAT05869) was isolated that encoded the	
CC	human LHR protein (AAR83050). LHR-IgG hybrids were constructed for use	
CC	in the targeting of therapeutic moieties to lymphoid tissue.	
XX		
SQ	Sequence 372 AA;	
QY	Query Match	100.0%; Score 215; DB 16; Length 372;
	Best Local Similarity	100.0%; Pred. No. 3e-14;
	Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 COPMSCGHGECVEIINNHCTNCDDGYGPOCOL 34	
	160 cqpwcsgthgevcvelimhtncdvgy99pqc1 193	
RESULT 26		
ID AAR98106	AAR98106 standard; Protein; 372 AA.	
XX		
AC AAR98106;		
XX		
DT 31-OCT-1996	(first entry)	
XX		
DE Human lymphocyte cell surface glycoprotein (HLHR).		

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..38
 FT /label= Signal region.
 FT Domain 39..155
 FT /label= Lectin domain.
 FT Domain 160..193
 FT /label= EGF domain.
 FT Binding-site 197..258
 FT /label= Complement binding repeat 1.
 FT Binding-site 259..317
 FT /label= Complement binding repeat 2.
 FT Domain 333..355
 FT /label= Transmembrane domain.
 FT Domain 356..372
 FT /label= Cytoplasmic domain.
 XX
 PN US514582-A.
 XX
 PD 07-MAY-1996.
 XX
 PF 23-FEB-1989; 89US-0315015.
 XX
 PR 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 DR WPI: 1996-238773/24.
 DR N-PSDB; AAR98106.
 XX
 PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
 PT binding site of a receptor fused to Ig constant region - useful for
 PT diagnosis and treatment e.g. of inflammation
 XX
 PS Example 2; Figure 1; 41pp; English.
 XX
 CC A hybrid immunoglobulin chain comprising the ligand binding site of
 CC a single transmembrane receptor without an active transmembrane
 CC region; fused at its C-terminus with the N-terminus of an
 CC immunoglobulin constant region. The receptor is not a member of the
 CC immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC the adhesion/targeting of a ligand binding partner (LBP) with the
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver
 CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery, improves aqueous solubility and removes
 CC potentially immunogenic epitopes.
 CC
 XX Sequence 372 AA;
 XX

OY 1 COWSCSGHCEVEITNNHTCMCDVGYGPOCOL 34
 ||||||||||||||||||||||||||||||||
 DB 160 cqwscsgheceveitnnhtcncdvgygppcql 193
 RESULT 27
 AAR98133
 ID AAR98133 standard; Protein; 372 AA.
 XX
 AC AAR98133;
 XX
 DT 01-NOV-1996 (first entry)
 XX
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
 XX
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 KW
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Region 1..37
 FT /label= Signal region.
 FT Domain 38..154
 FT /label= Lectin domain.
 FT Domain 159..192
 FT /label= EGF domain.
 FT Binding-site 196..258
 FT /label= Complement binding repeat 1.
 FT Binding-site 259..317
 FT /label= Complement binding repeat 2.
 FT Domain 333..355
 FT /label= Transmembrane domain.
 FT Domain 356..372
 FT /label= Cytoplasmic domain.
 XX
 PN US514582-A.
 XX
 PD 07-MAY-1996.
 XX
 PF 23-FEB-1989; 89US-0315015.
 XX
 PR 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 DR WPI: 1996-238773/24.
 XX
 PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
 PT binding site of a receptor fused to Ig constant region - useful for
 PT diagnosis and treatment e.g. of inflammation
 XX
 PS Disclosure: Page 19; 41pp; English.
 XX
 CC A hybrid immunoglobulin chain comprising the ligand binding site of
 CC a single transmembrane receptor without an active transmembrane
 CC region; fused at its C-terminus with the N-terminus of an
 CC immunoglobulin constant region. The receptor is not a member of the
 CC immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC the adhesion/targeting of a ligand binding partner (LBP) with the
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC treating graft rejection; inflammation; metastasis of Lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery, improves aqueous solubility and removes
 CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-P88135. This variant contains a 209-Aasn insertion.
 SQ Sequence 372 AA;

Query Match 100.0%; Score 215; DB 17; Length 372;
 Best Local Similarity 100.0%; Pred. No. 3e-14;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 COPWSCSGHGEVEIINHTCNCDDVGYGPQQL 34
 |||
 DB 159 cqpwsccghgeveiiinhhtcncddvgygpqql 192

RESULT 28
 AAW37781
 ID AAW37781 standard; Protein; 372 AA.
 XX
 AC AAW37781;
 XX
 DT 17-AUG-1998 (first entry)
 XX
 DE Homo sapiens lymphocyte homing receptor (LHR).
 XX
 KW lymphocyte homing receptor; LHR; HULHR; organ; graft; rejection;
 KW treatment; inflammatory disorders; rheumatoid arthritis;
 KW autoimmune diseases; lymphoma metastasis; control; lymphocyte;
 KW accumulation.
 XX
 OS Homo sapiens.

Location/Qualifiers
 Key 20..33
 FH Peptide /note= "potential signal sequence"
 FT /note= "potential N-linked glycosylation site"
 FT 60..62
 FT /note= "potential N-linked glycosylation site"
 FT 104..106
 FT /note= "potential N-linked glycosylation site"
 FT 177..179
 FT /note= "potential N-linked glycosylation site"
 FT 216..218
 FT /note= "potential N-linked glycosylation site"
 FT 232..234
 FT /note= "potential N-linked glycosylation site"
 FT 246..248
 FT /note= "potential N-linked glycosylation site"
 FT 271..273
 FT /note= "potential N-linked glycosylation site"
 FT 311..313
 FT /note= "potential N-linked glycosylation site"
 FT 335..357
 FT /note= "membrane anchoring domain/stop transfer"

US5714147-A.
 PD 03-FEB-1998.
 XX
 XX 23-FEB-1989; 89US-0315015.
 XX
 XX 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 19-DEC-1991; 91US-0808122.
 PR 08-DEC-1997; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.

PR 26-MAY-1995; 95US-0451848.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 DR WPI; 1998-129805/12.
 DR N-PSDB; AAV19012.
 XX
 PT Prevention of lymphocyte attachment to endothelial cells - using
 PT chimeric molecule comprising lymphocyte homing receptor and
 PT immunoglobulin constant region
 XX
 PS Disclosure; Fig 1; 43pp; English.
 XX
 CC The sequence is that of a human lymphocyte homing receptor
 CC (LHR) which may be used in the construction of a chimeric molecule
 CC comprising an LHR fused at its C terminus to the N terminus of an
 CC immunoglobulin constant region. This can be used for the prevention
 CC of lymphocyte attachment to endothelial cells. Such a method may
 CC be used for preventing organ or graft rejection, for treating
 CC inflammatory disorders, e.g. rheumatoid arthritis or other
 CC autoimmune diseases, for controlling lymphoma metastasis and
 CC for treating conditions in which there is an accumulation of
 CC lymphocytes.
 XX
 SQ Sequence 372 AA;

Query Match 100.0%; Score 215; DB 19; Length 372;
 Best Local Similarity 100.0%; Pred. No. 3e-14;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 COPWSCSGHGEVEIINHTCNCDDVGYGPQQL 34
 |||
 DB 160 cqpwsccghgeveiiinhhtcncddvgygpqql 193

RESULT 29
 AAW73264
 ID AAW73264 standard; Protein; 372 AA.
 XX
 AC AAW73264;
 XX
 DT 02-FEB-1999 (first entry)
 XX
 DE Human lymphocyte homing receptor.
 XX
 KW lymphocyte homing receptor; LHR; lymphocyte cell-surface glycoprotein;
 KW lymphocyte binding; endothelium; graft rejection; inflammation; therapy;
 KW arthritis; autoimmune disease; lymphoma metastasis;
 KW lymphocyte accumulation; human.
 XX
 OS Homo sapiens.

US5840844-A.
 PN 24-NOV-1998.
 XX
 PD 10-AUG-1995; 95US-0513278.
 PF
 XX 23-FEB-1989; 89US-0315015.
 PR 31-OCT-1991; 91US-0786149.
 PR 06-MAY-1993; 93US-0059029.
 PR 10-AUG-1995; 95US-0513278.

(GETH) GENENTECH INC.
 PA (REGC) UNIV CALIFORNIA.
 XX
 XX Lasky LA, Rosen SD, Slinger MS, Stachel SE;
 PI WPI; 1999-034122/03.
 XX
 DR N-PSDB; AAV08321.

XX Lymphocyte homing receptor polypeptides - useful for inhibiting
PT Lymphocyte binding to lymphoid endothelium
XX
PS Claim 1; Fig 1; 33pp; English.
XX
CC This sequence is the human lymphocyte homing receptor (LHR) of the
CC invention. LHR is a lymphocyte cell-surface glycoprotein that mediates
CC the binding of lymphocytes to the endothelium of lymphoid tissue. Soluble
CC LHR polypeptides, lacking signal peptide (amino acids 1-38),
CC transmembrane domain (amino acids 333-355) and cytoplasmic domain (amino
CC acids 356-372), can be used therapeutically to compete with the normal
CC binding of lymphocytes to lymphoid tissue and are especially useful for
CC organ or graft rejection treatment protocols, for treating inflammations
CC such as arthritis and other autoimmune diseases, for control of lymphoma
CC metastasis and for treating conditions involving lymphocyte accumulation.
CC LHR polypeptides can also be used in assays for LHR, anti-LHR antibodies
CC or competitive inhibitors of LHR activity, and for purifying anti-LHR
CC antibodies, and as immunogens for raising anti-LHR antibodies.
XX
SO Sequence 372 AA:

Query Match 100.0%; Score 215; DB 20; Length 372;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQWSCSGHGECEIINHTCNCMDVGYGPQCQL 34
Db 160 CQWSCSGHGECEIINHTCNCMDVGYGPQCQL 193

RESULT 30
AAR98131
ID AAR98131 standard; Protein; 374 AA.
XX
AC AAR98131;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HuLHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
KM
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..157
FT /label= Lectin domain.
FT Domain 162..195
FT /label= EGF domain.
FT Binding-site 199..260
FT /label= Complement binding repeat 1.
FT Binding-site 261..319
FT /label= Complement binding repeat 2.
FT Domain 335..357
FT /label= Transmembrane domain.
FT Domain 358..374
FT /label= Cytoplasmic domain.
XX
PN US5514582-A.
XX
PD 07-MAY-1996.
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI. 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PR binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure: Page 19; 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a 67-Glu-Ser-Ala insertion.
XX
SO Sequence 374 AA:

Query Match 100.0%; Score 215; DB 17; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQWSCSGHGECEIINHTCNCMDVGYGPQCQL 34
Db 162 CQWSCSGHGECEIINHTCNCMDVGYGPQCQL 195

RESULT 31
AAR98132
ID AAR98132 standard; Protein; 374 AA.
XX
AC AAR98132;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HuLHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
KM
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..157
FT /label= Lectin domain.
FT Domain 162..195
FT /label= EGF domain.
XX

FT	Binding-site	199..260
FT	/label= Complement binding repeat 1.	
FT	Binding-site	261..319.
FT	/label= Complement binding repeat 2.	
FT	Domain	335..357
FT	/label= Transmembrane domain.	
FT	Domain	358..374
FT	/label= Cytoplasmic domain.	
XX		
PN	US\$514582.A.	
XX		
PD	07-MAY-1996.	
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
PA	(GETH) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
XX	WPI; 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
XX		
XX	Disclosure; Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site, of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAR98106 are given in	
CC	AAR98109-R98135. This variant contains a 83-Gly-Thr insertion.	
XX		
SQ	Sequence 374 AA:	
OY	Query Match	100.0%; Score 215; DB 17; Length 374;
ID	Best Local Similarity	100.0%; Pred. No. 3.le-14;
DB	Matches 34; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	COPWSCGHGCEVEIINHTNCMDGYGYPOCOL	34
162	cqpwscgphgeveeiimhtncmdvgyygpgcql	195
RESULT 32		
AAR98134		
ID	AAR98134 standard; Protein; 374 AA.	
XX		
AC	AAR98134; `	
XX		
DT	01-NOV-1996 (first entry)	

XX	Human lymphocyte cell surface glycoprotein (HuLHR) variant.
DE	
XX	
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;
KM	diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KM	immunomodulator; cell adhesion; graft rejection; inflammation;
XX	metastasis.
OS	
XX	Homo sapiens.
PH	
Key	Location/Qualifiers
FT	Region
FT	1..37
FT	/label= Signal region.
FT	38..154
FT	/label= Lectin domain.
FT	159..192
FT	/label= EGF domain.
FT	196..260
FT	/label= Complement binding repeat 1.
FT	261..319
FT	/label= Complement binding repeat 2.
FT	335..357
FT	/label= Transmembrane domain.
FT	358..374
FT	/label= Cytoplasmic domain.
PN	
XX	US5514582-A.
PD	
XX	07-MAY-1996.
XX	
PF	23-FEB-1989; 89US-0315015.
XX	
PR	22-NOV-1989; 89US-0440625.
PR	23-FEB-1989; 89US-0315015.
PR	16-DEC-1991; 91US-0808122.
PR	08-DEC-1992; 92US-0986931.
PR	21-JAN-1994; 94US-0185610.
XX	
PA	(GERTH) GENENTECH INC.
P1	
XX	Capon DJ, Lasky LA;
DR	
XX	WPI, 1996-238773/24.
XX	
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT	binding site of a receptor fused to Ig constant region - useful for
PT	diagnosis and treatment e.g. of inflammation
XX	
PS	Disclosure; Page 19; 41pp; English.
XX	
CC	A hybrid immunoglobulin chain comprising the ligand binding site of
CC	a single transmembrane receptor without an active transmembrane
CC	region; fused at its C-terminus with the N-terminus of an
CC	immunoglobulin constant region. The receptor is not a member of the
CC	immunoglobulin super family, nor a multiple subunit polypeptide
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines
CC	the adhesion/targeting of a ligand binding partner (LBP) with the
CC	effector functions of immunoglobulin and can bind to and/or activate
CC	more than one ligand. It can be used diagnostically for the in
CC	vitro assay of LBP and their targets; or therapeutically to deliver
CC	LBP such as toxins, enzymes, growth factors to particular cells.
CC	Typical applications are as antiviral, neuromodulating and
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC	treating graft rejection, inflammation; metastasis of lymphoma etc.).
CC	The immunoglobulin component increases plasma half life and
CC	facilitates purification while deletion of the transmembrane region
CC	accelerates recovery, improves aqueous solubility and removes
CC	cell surface immunogenic epitopes. Variants of the human lymphocyte
CC	cell surface glycoprotein described in AAP8106 are given in
CC	AA98109-R98135. This variant contains a 241-Val-Glu-Asn insertion
XX	
Sequence	374 AA;
50	

Query Match 100.0%; Score 215; DB 17; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQWSCSGHGECEIINHTCNCVDYGYGPQCQL 34
Db 159 cqwscsghgeceiinhtcncdvygygpqcql 192

RESULT 33
AAR98135
ID AAR98135 standard; Protein; 374 AA.
AC AAR98135;
XX
XX 01-NOV-1996 (first entry)
DT
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
DE
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW Immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 1..37
FT /label= Signal region.
FT 38..154
FT /label= Lectin domain.
FT 159..192
FT Domain /label= EGF domain.
FT 196..257
FT Binding-site /label= Complement binding repeat 1.
FT 258..319
FT Binding-site /label= Complement binding repeat 2.
FT 335..357
FT Domain /label= Transmembrane domain.
FT 358..374
FT Domain /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
PD
XX
XX 23-FEB-1989; 89US-0315015.
PE
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX
XX (GETH) GENENTECH INC.
PA
XX
XX Capon DJ, Lasky LA;
PI
XX
XX WPI: 1996-238773/24.
DR
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX
XX Disclosure; Page 19; 41pp; English.
PS
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines

CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets, or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a 292-Tyr-Tyr-Tyr Insertion.
XX
XX
XX Sequence 374 AA;
SO

Query Match 100.0%; Score 215; DB 17; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQWSCSGHGECEIINHTCNCVDYGYGPQCQL 34
Db 159 cqwscsghgeceiinhtcncdvygygpqcql 192

RESULT 34
AAR98115
ID AAR98115 standard; Protein; 371 AA.
AC AAR98115;
XX
XX 01-NOV-1996 (first entry)
DT
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
DE
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW Immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 1..37
FT /label= Signal region.
FT 38..154
FT /label= Lectin domain.
FT 159..192
FT Domain /label= EGF domain.
FT 196..257
FT Binding-site /label= Complement binding repeat 1.
FT 258..316
FT Binding-site /label= Complement binding repeat 2.
FT 332..354
FT Domain /label= Transmembrane domain.
FT 355..371
FT Domain /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
PD
XX
XX 23-FEB-1989; 89US-0315015.
PE
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX
XX (GETH) GENENTECH INC.
PA

```

XX  Capon DJ, Lasky LA;
PI
XX
XX  WPI, 1996-238773/24.
DR
XX
XX  Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX  Disclosure; Page 19; 41pp; English.
XX
XX  A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an Ile174Leu substitution.
XX
XX  Sequence 371 AA;
SQ
XX
XX  Query Match 99.1%; Score 213; DB 17; Length 371;
Best Local Similarity 97.1%; Pred. No. 4.8e-14;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 COPWSCSGHGECEIINNHTCNCDVGYGPOCOL 34
Db 159 cqpwschgcevelinhhtcndvgygpcqql 192
RESULT 35
AAR98116
ID AAR98116 standard; Protein; 371 AA.
XX
XX  AAR98116;
AC
XX
XX  01-NOV-1996 (first entry)
DT
XX
XX  Human lymphocyte cell surface glycoprotein (HULHR) variant.
DE
XX
XX  Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
XX  Homo sapiens.
OS
XX
XX  Key Location/Qualifiers
FH 1..37
FT /label= Signal region.
FT 38..154
FT /label= Lectin domain.
FT 159..192
FT /label= EGF domain.
FT 196..257
FT /label= Complement binding repeat 1.
FT 258..316
FT /label= Complement binding repeat 2.
FT 332..354
FT Domain

```

```

FT  /label= Transmembrane domain.
FT 355..371
FT /label= Cytoplasmic domain.
XX
XX  US5514582-A.
XX
XX  07-MAY-1996.
XX
XX  23-FEB-1989; 89US-0315015.
XX
XX  22-NOV-1989; 89US-0440625.
XX  23-FEB-1989; 89US-0315015.
XX  16-DEC-1991; 91US-0808122.
XX  08-DEC-1992; 92US-0986931.
XX  21-JAN-1994; 94US-0185670.
XX
XX  (GETH ) GENENTECH INC.
XX
XX  Capon DJ, Lasky LA;
PI
XX
XX  WPI, 1996-238773/24.
DR
XX
XX  Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX  Disclosure; Page 19; 41pp; English.
XX
XX  A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an Asn181Gln substitution.
XX
XX  Sequence 371 AA;
SQ
XX
XX  Query Match 97.2%; Score 209; DB 17; Length 371;
Best Local Similarity 97.1%; Pred. No. 1.2e-13;
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 COPWSCSGHGECEIINNHTCNCDVGYGPOCOL 34
Db 159 cqpwschgcevelinhhtcndvgygpcqql 192
RESULT 36
AAR98114
ID AAR98114 standard; Protein; 371 AA.
XX
XX  AAR98114;
AC
XX
XX  01-NOV-1996 (first entry)
DT
XX
XX  Human lymphocyte cell surface glycoprotein (HULHR) variant.
DE
XX
XX  Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW

```

KW Immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..37
 FT Domain /label= Signal region.
 FT Domain 38..154
 FT Domain /label= Lectin domain.
 FT Domain 159..192
 FT Domain /label= EGF domain.
 FT Binding-site 196..257
 FT Binding-site /label= Complement binding repeat 1.
 FT Binding-site 258..316
 FT Binding-site /label= Complement binding repeat 2.
 FT Domain 332..354
 FT Domain /label= Transmembrane domain.
 FT Domain 355..371
 FT Domain /label= Cytoplasmic domain.
 XX
 PN US5514582-A.
 XX
 PD 07-MAY-1996.
 XX
 PF 23-FEB-1989; 89US-0315015.
 XX
 PR 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 DR WPI: 1996-238773/24.
 XX
 XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
 PT binding site of a receptor fused to Ig constant region - useful for
 PT diagnosis and treatment e.g. of inflammation
 XX
 PS Disclosure; Page 19; 41pp; English.
 XX
 XX A hybrid immunoglobulin chain comprising the ligand binding site of
 CC a single transmembrane receptor without an active transmembrane
 CC region; fused at its C-terminus with the N-terminus of an
 CC immunoglobulin constant region. The receptor is not a member of the
 CC immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC the adhesion/targeting of a ligand binding partner (LBP) with the
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver
 CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery. Improves aqueous solubility and removes
 CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-R98135. This variant contains an His168Gln substitution.
 CC
 XX Sequence 371 AA;
 SQ

Query Match 96.3%; Score 207; DB 17; Length 371;
 Best Local Similarity 97.1%; Pred. No. 1.9e-13;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 COPMSCSGHECVETIINNNHCNDVGYGPOCOL 34
 ||||||| ||||||| ||||||| ||||||| |||||||
 Db 159 cqpmscsgqgecvellnhntcndvgyy9pqcql 192

RESULT 37
 AA57762
 ID AA57762 standard; protein. 196 AA.
 AC
 AC AA57762;
 DT 17-MAR-2000 (first entry)
 XX
 DE Human L-selectin amino acid sequence.
 XX
 KW E-selectin; L-selectin; P-selectin; bifunctional molecule; regulation;
 KW binding; detection; cellular proliferation; cytostatic; cell adhesion;
 KW antiinflammatory; immunosuppressant; immunostimulator; cell migration;
 KW tumour; neoplasia; cancer; inflammation.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 155
 FT /note= "Unclear in specification"
 XX
 XX W09961033-A1.
 XX
 PD 02-DEC-1999.
 XX
 PE 21-MAY-1999; 99WO-US11300.
 XX
 PR 22-MAY-1998; 98US-0086442.
 XX
 PA (UYHO-) UNIV HOUSTON.
 XX
 PI Freidman J;
 XX
 DR WPI: 2000-072543/06.
 XX
 XX Novel bifunctional compounds, used for treating e.g. cancers,
 PT inflammation or transplant rejection or for tissue repair or
 PT regeneration -
 XX
 PS Example 2; Fig 3; 60pp; English.
 XX
 XX The present invention describes a bifunctional compound for the
 CC regulation of cellular proliferation. The compound comprises, a
 CC cell-adhesion oligosaccharide (CAOS), a linker group attached to the
 CC CAOS by the reducing end of the CAOS or by a primary hydroxyl group,
 CC and a nucleotide cyclic-3'-5' monophosphate attached to the linker
 CC group through a heterocyclic base. The bifunctional compounds can be
 CC used for regulating cell proliferation, cell adhesion and cell
 CC migration. Methods from the present invention can also be used to
 CC identify agonists or antagonists of such activities. The bifunctional
 CC compounds can be used for treating e.g. tumours, neoplasia, cancer or
 CC inflammation. They can also be used to serve as immunosuppressants in
 CC certain applications such as organ or tissue transplantation, as
 CC immunostimulators and to provide immunological activity for disease
 CC such as AIDS. They can also have important uses in the control and
 CC generation and growth for new tissue. This can be useful in the repair
 CC and regeneration of tissue as well as in the transplantation or
 CC introduction of artificial tissue into the body. This can have
 CC beneficial effects in the treatment of burns or other replacement
 CC therapies requiring increased cell-cell adhesion, cell migration and/or
 CC cellular growth. The present sequence represents a human L-selectin
 CC amino acid sequence given in the exemplification of the present
 CC invention.
 XX
 XX Sequence 196 AA;
 SQ

Query Match 95.3%; Score 205; DB 21; Length 196;

Best Local Similarity 97.0%; Pred. No. 1.8e-13;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINHTCNCDVGYGPQCQ 33
Db 160 cqpwsqsgnhecveilnnyncncdvgygpqcq 192

RESULT 38
AAR22551

ID AAR22551 standard; Protein; 363 AA.

AC AAR22551;

DT 21-MAY-1992 (first entry)

DE T lymphocyte-specific Leu8 Antigen minor form.

KW Rapid immunoselection cloning technique; cell surface antigen;
homing receptor; antigen-presenting cells.

OS Homo sapiens.

Key Location/Qualifiers

Modified-site 113..115
/label= N-linked_glycosylation

Modified-site 157..159
/label= N-linked_glycosylation

Modified-site 230..232
/label= N-linked_glycosylation

Modified-site 269..271
/note= "putative"

Modified-site 285..287
/label= N-linked_glycosylation

Modified-site 299..301
/label= N-linked_glycosylation

Modified-site 324..326
/label= N-linked_glycosylation

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

Modified-site
/note= "putative"

CC See AAQ21184 for the larger insert and AAR20815 for the major form of
CC the Leu8 antigen that it encodes.
XX

Sequence 363 AA;

Query Match 95.3%; Score 205; DB 13; Length 363;
Best Local Similarity 97.0%; Pred. No. 3e-13;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINHTCNCDVGYGPQCQ 33
Db 173 cqpwsqsgnhecveilnnyncncdvgygpqcq 205

RESULT 39
AAR91443

ID AAR91443 standard; Protein; 363 AA.

AC AAR91443;

DT 31-OCT-1996 (first entry)

DE Human Leu8 antigen.

KW Cell surface antigen; cloning; immunoselection; immunotherapy;
therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;
antiinflammatory.

OS Homo sapiens.

Key Location/Qualifiers

Modified-site 113..115
/label= Glycosylation_site

Modified-site 156..158
/label= Glycosylation_site

Modified-site 229..231
/label= Glycosylation_site

Modified-site 269..271
/label= Glycosylation_site

Modified-site 285..287
/label= Glycosylation_site

Modified-site 299..301
/label= Glycosylation_site

Modified-site 324..326
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

Modified-site
/label= Glycosylation_site

CC cloning method. The longer insert (AA114723) contained 2,350
CC residues, while the shorter lacked 436 internal residues. A major
CC transcript of 2.4 kb was present in peripheral blood mononuclear
CC cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,
CC and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukemic
CC T-cell lines. Leu8 antigens can be obtd. for diagnostic and
CC therapeutic use. The presence or absence of Leu8 on CD4+ T-cells
CC identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.
CC Soluble forms of Leu8 can act as antiinflammatory agents by reducing
CC lymphocyte migration.
CC
SQ Sequence 363 AA;

Query Match 95.3%; Score 205; DB 17; Length 363;
Best Local Similarity 97.0%; Pred. No. 3e-13;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECVEITNNHTCNCDDVGYGPQCO 33
DB 173 cqpwsctghgcevelnnhtcncddvgyygpqco 205

RESULT 40
AA96184
ID AA96184 standard; Protein: 363 AA.
XX
AC AA96184;
XX
DT 19-DEC-2000 (first entry)
XX
DE Human T-cell specific Leu8 antigen.
XX
KW Leu8; cell surface antigen; human; immunoselection; panning;
KW immunodiagnosis; diagnosis; immunotherapy; gene therapy;
KW immune disorder; infection; asthma; immune-complex disease;
KW amyloidosis; multiple sclerosis; inflammation; antiinflammatory.
XX
OS Homo sapiens.
XX
PN US611093-A.
XX
PD 29-AUG-2000.
XX
PE 28-OCT-1998; 98US-0181612.
XX
PR 01-DEC-1992; 92US-0983647.
PR 25-FEB-1988; 88US-0160416.
PR 13-JUL-1989; 89US-0379076.
PR 23-MAR-1990; 90US-0498809.
PR 13-JUL-1990; 90US-0553759.
XX
PA (GENO) GEN HOSPITAL CORP.
XX
PI Stamenkovic I, Seed B;
XX
DR WPI: 2000-586382/55.
DR N-PSDB; AAA50632.
XX
XX Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
PT useful for immunodiagnosis and immunotherapy of immune-mediated
PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
PT diseases -
XX
XX Example 14; Column 69-72; 75pp; English.
XX
CC The present sequence is that of a T-cell specific Leu8 antigen, as
CC predicted from 1 of 2 clones (see AA50632) isolated from a human
CC T-cell library by the method of the invention. A longer Leu8
CC antigen is given in AA96184. The method, designed to isolate cell
CC surface antigen (CSA) cDNAs, is based upon transient expression of
CC a CSA in eukaryotic cells and physical selection of cells expressing
CC the antigen by adhesion to (panning on) an antibody-coated substrate

CC such as a culture dish. CSA nucleic acids isolated by the method of
CC the invention, and the proteins they encode, are useful for
CC immunodiagnostic and immunotherapeutic applications, including the
CC diagnosis and treatment of immune-mediated infections, diseases, and
CC disorders in animals, including humans. These disorders include
CC asthma, immune-complex disease, amyloidosis, parasitic diseases or
CC multiple sclerosis. The ability to interfere with the binding of
CC Leu8-T-cells to antigen presenting cells, or the ability to cause
CC such binding to occur on surfaces other than lymphocyte cells, can
CC be useful in diagnostics and therapy. The level of activated Leu8-
CC T-cells relative to resting Leu8+ cells could serve as a measure of
CC immune response to a particular antigen. Modification of the
CC specificity of the extracellular domain of Leu8, which mediates
CC adhesion to specific endothelial cells of lymph nodes, could serve
CC to regulate the homing potential of resting T cells. Soluble forms
CC of Leu8 could act as antiinflammatory agents by reducing lymphocyte
CC migration.
CC
SQ Sequence 363 AA;

Query Match 95.3%; Score 205; DB 21; Length 363;
Best Local Similarity 97.0%; Pred. No. 3e-13;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECVEITNNHTCNCDDVGYGPQCO 33
DB 173 cqpwsctghgcevelnnhtcncddvgyygpqco 205

Search completed: September 7, 2002, 10:14:07
Job time: 265 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:20 ; Search time 61.04 Seconds

(without alignments)
374.104 Million cell updates/sec

Title: US-09-119-209-2_COPY_197_328

Perfect score: 733

Sequence: 1 CEPLAEPLGTMDCTHFGN.....WSNPSPICQKLDKSFMIKE 132

Scoring table:

BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_oranelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_ricent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	99.2	385	4 Q9UJ43	Q9UJ43 homo sapien
2	599	81.7	376	6 Q28629	Q28629 oryctolagus
3	535	73.0	372	11 Q63762	Q63762 rattus norv
4	318	43.4	646	6 Q29097	Q29097 sus scrofa
5	315	43.0	740	4 Q95508	Q95508 homo sapien
6	310	42.3	610	6 Q951G1	Q951G1 equus cabal
7	309	42.2	616	4 Q95509	Q95509 homo sapien
8	307	41.9	740	4 Q95507	Q95507 homo sapien
9	306	41.7	754	6 Q28290	Q28290 canis fami1
10	300	40.9	609	6 Q9G1F0	Q9G1F0 canis fami1
11	291	39.7	649	6 Q28657	Q28657 oryctolagus
12	290.5	39.6	485	6 Q951G3	Q951G3 odocoileus
13	275.5	37.6	484	6 Q951G2	Q951G2 ovis aries
14	245	33.4	482	6 Q28982	Q28982 sus scrofa
15	192	26.2	3567	11 Q9ES77	Q9ES77 mus musculu
16	183	25.0	866	5 Q917E3	Q917E3 drosophila

17	183	25.0	933	5 Q917E4	Q917E4 drosophila
18	183	25.0	958	5 Q9V560	Q9V560 drosophila
19	173	23.6	1172	4 Q9NU87	Q9NU87 homo sapien
20	173	23.6	1653	5 Q9VITU9	Q9VITU9 drosophila
21	172.5	23.5	481	4 Q9H284	Q9H284 homo sapien
22	169	23.1	1236	11 Q91YB6	Q91YB6 rattus norv
23	168.5	23.0	1316	4 Q96J07	Q96J07 homo sapien
24	167.5	22.9	2014	6 Q29530	Q29530 pan troglod
25	163.5	22.3	2039	4 Q16745	Q16745 homo sapien
26	163.5	22.3	2489	4 Q16744	Q16744 homo sapien
27	160	21.8	3389	4 Q96C09	Q96C09 homo sapien
28	160	21.8	3508	4 Q96RM4	Q96RM4 homo sapien
29	156.5	21.4	390	11 Q921P0	Q921P0 mus musculu
30	155	21.1	533	11 Q08569	Q08569 cavia porce
31	155	21.1	555	11 Q99JA1	Q99JA1 cavia porce
32	154.5	21.1	222	6 Q19122	Q19122 callimico g
33	153	20.9	1911	6 Q29528	Q29528 papio hamad
34	153	20.9	3564	11 Q923L3	Q923L3 mus musculu
35	152.5	20.8	363	6 Q02839	Q02839 sus scrofa
36	152.5	20.8	395	12 Q9J2M6	Q9J2M6 macaca mula
37	151.5	20.7	354	5 Q9U6I1	Q9U6I1 drosophila
38	151.5	20.7	410	4 Q9H4R2	Q9H4R2 homo sapien
39	151.5	20.7	465	4 Q60687	Q60687 homo sapien
40	150.5	20.5	560	5 Q22328	Q22328 caenorhabdl
41	150.5	20.5	1032	11 Q9DC83	Q9DC83 mus musculu
42	149.5	20.4	378	6 Q62837	Q62837 saguinus oe
43	149	20.3	315	6 Q28770	Q28770 papio cynoc
44	149	20.3	522	6 Q28769	Q28769 papio cynoc
45	148.5	20.3	559	4 Q9UQV2	Q9UQV2 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	385 AA.
Q9UJ43	Q9UJ43	Q9UJ43		
AC	Q9UJ43	Q9UJ43		
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	L-SELECTIN PRECURSOR.			
GN	L-SELECTIN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
NCBI_Taxid=9606;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);			
RA	Fieger C.B.;			
RL	Submitted (Sep-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);			
RA	Fieger C.B.;			
RL	Thesis (1998), Freie Universitaet Berlin, Fachbereich Chemie.			
DR	EMBL; AJ246000; CAB55488.1; -.			
DR	HSSP; P14151; 1KB.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001304; lectin.C.			
DR	InterPro; IPR002396; Selectin.			
DR	InterPro; IPR000436; Sush1_SCR_CCP.			
DR	Pfam; PF00008; EGF_1.			
DR	Pfam; PF00059; lectin.C_1.			
DR	Pfam; PF00084; sush1_2.			
DR	PRINTS; PR00343; SELECTIN.			
DR	SMART; SM00032; CCP_2.			
DR	SMART; SM00034; CLECT_1.			
DR	SMART; SM00181; EGF_1.			
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.			
DR	PROSITE; PS00041; C_TYPE_LECTIN_2; 1.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.			

DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.
FT SIGNAL 1 51 POTENTIAL.
FT CHAIN 52 385 L-SELECTIN
SQ SEQUENCE 385 AA; 43617 MW; 1205F691BA638EP1 CRC64;

Query Match 99.2%; Score 727; DB 4; Length 385;
Best Local Similarity 99.2%; Pred. No. 2,1e-78;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAELGTMDCTHPGNFSFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60
DB 210 CEPLAELGTMDCTHPGNFSFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 269
QY 61 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGEETLIGKKTKICSSGIWNSPSPIC 120
DB 270 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGEETLIGKKTKICSSGIWNSPSPIC 329
QY 121 QKLDKSFMSIKE 132
DB 330 QKLDKSFMSIKE 341

RESULT 2
ID Q28629 PRELIMINARY; PRT; 376 AA.
AC Q28629;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE L-SELECTIN PRECURSOR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY CORTEX;
RA Olan J., Marks R.M.;
RT "CDNA for rabbit L-selectin.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U26535; AAA67896.1; -
DR HSSP; P14151; 1KJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin.c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sushl_SCR_CCP.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; lectin_c; 1.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 2.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 1 385 L-SELECTIN.
SQ SEQUENCE 376 AA; 42346 MW; 59F6AD530F490947 CRC64;

Query Match 81.7%; Score 599; DB 6; Length 376;
Best Local Similarity 80.3%; Pred. No. 3.8e-63;
Matches 106; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

QY 1 CEPLAELGTMDCTHPGNFSFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60
DB 197 CEPLAELGTMDCTHPGNFSFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 256

QY 61 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGEETLIGKKTKICSSGIWNSPSPIC 120
DB 257 IQCEPLAPDGLGIDCHPRAVFGFTSTCTFCSSGALLIMKKTKYCGSSGIWNSPSPIC 316
QY 121 QKLDKSFMSIKE 132
DB 317 QKLDKSFMSIKE 328

RESULT 3
ID Q63762 PRELIMINARY; PRT; 372 AA.
AC Q63762;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE LYMPOCYTE MEMBRANE PROTEIN A.11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RC MEDLINE=95369821; PubMed=7543874;
RX Sackstein R., Meng L., Xu X.M., Chin Y.H.;
RT "Evidence of post-transcriptional regulation of L-selectin gene
RT expression in rat lymphoid cells.";
RL Immunology 85:198-204(1995).
DR EMBL; S79523; AAC60710.2; -
DR HSSP; P14151; 1KJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin.c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sushl_SCR_CCP.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushl; 2.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 2.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 372 AA; 42471 MW; 7DFD125610DD6E4A CRC64;

Query Match 73.0%; Score 535; DB 11; Length 372;
Best Local Similarity 71.2%; Pred. No. 1.6e-55;
Matches 94; Conservative 15; Mismatches 23; Indels 0; Gaps 0;

QY 1 CEPLAELGTMDCTHPGNFSFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60
DB 197 CEPLAELGTMDCTHPGNFSFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 256
QY 61 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGEETLIGKKTKICSSGIWNSPSPIC 120
DB 257 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGEETLIGKKTKYCGSSGIWNSPSPIC 316
QY 121 QKLDKSFMSIKE 132
DB 317 QKLDKSFMSIKE 328

RESULT 4
ID Q29097 PRELIMINARY; PRT; 646 AA.
AC Q29097;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE P-SELECTIN PRECURSOR.
 OS Sus scrofa (pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 NC NCBITaxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-AORTA.
 RA Rollins S.A., Johnson K.K., Birks C.W., Matis L.A., Rother R.P.;
 RL Submitted (Oct-1995) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-AORTA;
 RX MEDLINE-20171534; PubMed-10706724;
 RA Stocker C.J., Sugars K.L., Harari O.A., Landis R.C., Morley B.J.,
 RA Haskard D.O.;
 RT "TNF-alpha, IL-4, and IFN-gamma regulate differential expression of P-
 RT and E-selectin expression by porcine aortic endothelial cells";
 RL J. Immunol. 164:3309-3315(2000).
 DR EMBL: L39075; AAA79007.1;
 DR EMBL: AF163766; AA43272.1;
 DR HSSP: P16109; IFSB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001304; lectin_c.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sush1; 6.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP; 6.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF; 1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 1.
 KW EGF-like domain; Glycoprotein; Lectin; Selectin; signal.
 FT SIGNAL
 FT CHAIN 1 41 POTENTIAL.
 FT CHAIN 42 646 P-SELECTIN.
 SQ SEQUENCE 646 AA; 71127 MW; 3863FA4FE09F0BB6 CRC64;

Query Match 43.4%; Score 318; DB 6; Length 646;
 Best Local Similarity 42.1%; Pred. No. 2.2e-29;
 Matches 51; Conservative 22; Mismatches 48; Indels 0; Gaps 0;
 QY 1 CEPLAPELGTMDCTHPNFSSSQCASFSGTNTLTGIEETGCGPFGNWSPEPTCOV 60
 DB 324 CEPLSPVAGSMDCPSSSAFQNTSCSRCAKGLTRADYVRCNSLQGTAPAPVCA 383
 QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFTCSGTELGKTKTCSSGIMSNSPTC 120
 DB 384 LQCODLPAPKQVNCVSHFAGFRVQSTGCTCDEGSLVGSVYLQCLFTGMSAPAPPC 443
 QY 121 Q 121
 DB 444 Q 444
 RESULT 5
 ID 095508 PRELIMINARY; PRT; 740 AA.
 AC 095508;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DJ780M13.1.2 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN
 DE CD62, GMP140)) (ISOFORM 2) (FRAGMENT).
 GN SELP.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBITaxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Howden P.;
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL022146; CA18143.1;
 DR HSSP: P16109; IKUD.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001304; lectin_c.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR Pfam: PF00008; EGF; 1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sush1; 8.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP; 8.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF; 1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 1.
 KW EGF-like domain; Glycoprotein.
 FT NON_TER 740
 FT 740
 SQ SEQUENCE 740 AA; 81390 MW; 1D2E35E6D93745CE CRC64;

Query Match 43.0%; Score 315; DB 4; Length 740;
 Best Local Similarity 43.9%; Pred. No. 6e-29;
 Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;
 QY 1 CEPLAPELGTMDCTHPNFSSSQCASFSGTNTLTGIEETGCGPFGNWSPEPTCOV 60
 DB 262 COHLEAPSEGTMDCTHPNFSSSQCASFSGTNTLTGIEETGCGPFGNWSPEPTCOV 321
 QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFTCSGTELGKTKTCSSGIMSNSPTC 120
 DB 322 IOCEPLSPVAGSMDCPSSSAFQNTSCSRCAKGLTRADYVRCNSLQGTAPAPVCA 381
 QY 121 Q 123
 DB 382 QAL 384

RESULT 6
 ID 0951G1 PRELIMINARY; PRT; 610 AA.
 AC 0951G1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE E-SELECTIN.
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 NC NCBITaxid=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21421234; PubMed-11529941;
 RA Hedges J.F., Demula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,
 RA MacLachlan N.J.;
 RT "Characterization of equine E-selectin";
 RL Immunology 103:498-504(2001).
 DR EMBL: AF307972; AA48712.1;
 KW Lectin; Selectin.
 SQ SEQUENCE 610 AA; 66191 MW; F9D3DD12C445382 CRC64;

Query Match 42.3%; Score 310; DB 6; Length 610;
 Best Local Similarity 44.6%; Pred. No. 1.9e-28;
 Matches 54; Conservative 16; Mismatches 51; Indels 0; Gaps 0;

AC	095507;	(TrEMBLrel. 10, Created)
DT	01-MAY-1999	(TrEMBLrel. 10, Last sequence update)
DR	EMBL; AL022146; CAA18142.1; -	
DR	HSSP; P16109; 1KJD.	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR001304; lectin_c.	
DR	InterPro; IPR002396; Selectin.	
DR	InterPro; IPR000436; Sushi_SCR_CCP.	
DR	Pfam; PF00008; EGF_1.	
DR	Pfam; PF00059; lectin_c; 1.	
DR	Pfam; PF00084; sushi; 8.	
DR	PRINTS; PR00343; SELECTIN.	
DR	SMART; SM00032; CCP; 8.	
DR	SMART; SM00034; CLECT; 1.	
DR	SMART; SM00181; EGF; 1.	
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.	
DR	PROSITE; PSS0041; C_TYPE_LECTIN_2; 1.	
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.	
DR	PROSITE; PS01186; EGF_2; 1.	
KW	EGF-like domain; Glycoprotein.	
FT	NON PER 740	
SEQUENCE	740 AA; 81095 MW; 3BSF70A45B1A3CD4 CRC64;	
Query Match	41.9%; Score 307; DB 4; Length 740;	
Best Local Similarity	40.7%; Pred. No. 5,4e-28;	
Matches	50; Conservative 24; Mismatches 49; Indels 0; Gaps 0;	
QY	1 CEPLAEELCTMDCTHPGNGSEFSQSACSEGNTLTGIETTCGGPGMWSPEPTCOV 60	
DH	200 CGEELTPEOHVLAMNCSDRLNFSNSQCSFYCHTDGYGVNGPSKECLIASGTWNPPQCLA 259	
QY	61 IOCEPIISAPDLTGIMNSDHLASFSFTSACTFCISCGTELLIKKKTIIESSGIMSNSPIC 120	
DH	260 AQCPPLKIPRKGMTLHLSAKAFQHOSSCSFSGCEGFALVEPVQCTASGVWTAPAVC 319	
QY	121 QKL 123	
DH	320 KAI 322	
RESULT	9	
ID	Q28290	PRELIMINARY; PRT; 754 AA.
AC	Q28290;	
DT	01-NOV-1996	(TrEMBLrel. 01, Created)
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)
DE	CELL ADHESION MOLECULE PRECURSOR (FRAGMENT).	
GN	GMP140.	
OS	Canis familiaris (Dog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.	
OX	NCBI_TaxID=9615;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=SPLEN.	
RA	Manning A.M., Sanders W.E.Jr., Kukielka G.L., Dore M.,	
RA	Rosenblom C.L., Hawkins H.L., Michael L.H., Entman M.L., Smith C.W.,	
RA	Baudet A.L., Anderson D.C.;	

"Molecular cloning of canine GMP140 and studies of expression in a
RT model of myocardial ischemia/reperfusion."
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; M88170; AAA63789.1; -.
DR HSSP; P16109; IFSB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_C.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00059; lectin_C; 1.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 8.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C-type-LECTIN_1; 1.
DR PROSITE; PS50041; C-type-LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein; Signal.
FT NON_TER 1
FT SIGNAL <1 13 POTENTIAL.
FT CHAIN 14 754 CELL ADHESION MOLECULE.
SQ SEQUENCE 754 AA; 82303 MM; F0438BEAA521E773 CRC64;

Query Match 41.7%; Score 306; DB 6; Length 754;
Best Local Similarity 40.5%; Pred. No. 7.3e-28;
Matches 49; Conservative 24; Mismatches 48; Indels 0; Gaps 0;

OY 1 CEPLAPELLGTMDCTHPGNFSSSOCAFCSEGTNLGIEPTTGGPFGMSSPPTCOY 60
DB 185 CGDFLPQHVLMNCNHPGNFSSSECTFYCTEGTELGPSPLELASTGTWTKPRCA 244
OY 61 IQCEPLASPDIGIMNCNHPPLASFSFTSACTFYCSGTELLIGKKTICSSGIMSNPSPIC 120
DB 245 TQCPPLKTPREGSNMCLHVSFAFYQSSCHFCCEGFPALVGEVYQCTASGMWTAAPVC 304
OY 121 Q 121
DB 305 E 305

RESULT 10
O9GLF0 PRELIMINARY; PRT; 609 AA.
AC O9GLF0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE E-SELECTIN.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOTHELIAL CELLS;
RA Zhang L., Shi Y., Wu H., Zhang G.;
RT "Cloning and sequencing of beagle E-selectin genomic DNA and
RT comparison with other species."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF287257; AAG10039.1; -.
DR HSSP; P16581; IFSB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_C.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00059; lectin_C; 1.
DR Pfam; PF00084; sush1; 6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 6.
DR SMART; SM00034; CLECT; 1.

DR SMART; SM00181; EGF; 4.
DR SMART; SM00001; EGF-like; 1.
DR PROSITE; PS50041; C-type-LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW lectin; Selectin.
SQ SEQUENCE 609 AA; 66073 MM; 41E62D1FAD23881F CRC64;

Query Match 40.9%; Score 300; DB 6; Length 609;
Best Local Similarity 43.0%; Pred. No. 2.9e-27;
Matches 52; Conservative 18; Mismatches 51; Indels 0; Gaps 0;

OY 1 CEPLAPELLGTMDCTHPGNFSSSOCAFCSEGTNLGIEPTTGGPFGMSSPPTCOY 60
DB 181 COAENPHEGSLVCTHPLGTFSTYNSCFYSCDKGLPSPSTETVQCTSTGEMASPPACNV 240
OY 61 IQCEPLASPDIGIMNCNHPPLASFSFTSACTFYCSGTELLIGKKTICSSGIMSNPSPIC 120
DB 241 VECASALTNPCHGVMDCLDSSGNFPNNMCTFCEGFPALMGPKRLQCTSSGMWDRKRPIC 300
OY 121 Q 121
DB 301 K 301

RESULT 11
Q28657
ID Q28657 PRELIMINARY; PRT; 649 AA.
AC Q28657;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P-SELECTIN.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Vora D.K., Fang Z., Liya S.M., Parhami F., Watson A.D., Drake T.A.,
RA Territo M.C., Berliner J.A.;
RT "Induction of P-selectin by KM-LDL and its role in human
RT atherosclerosis."
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Warden C.H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39446; AAR81385.1; -.
DR HSSP; P16109; IFSB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_C.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00059; EGF; 1.
DR Pfam; PF00059; lectin_C; 1.
DR Pfam; PF00084; sush1; 6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 6.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C-type-LECTIN_1; 1.
DR PROSITE; PS50041; C-type-LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain; Glycoprotein; lectin; Selectin.
SQ SEQUENCE 649 AA; 71755 MM; ECCD8C847B84BC31 CRC64;

Query Match 39.7%; Score 291; DB 6; Length 649;
Best Local Similarity 40.0%; Pred. No. 3.7e-26;
Matches 48; Conservative 25; Mismatches 47; Indels 0; Gaps 0;

QY 4 LEAPLGTMDCTHHPGNFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 63
 Db 203 LDLPQHVMNCHPLGNFSFNHSCSFHCADYALNGPSELCLASGIMTNPQCVAAQC 262
 QY 64 EPLAPDGLGIMNCHPLASFSFTSACTFTCSGTELLIGKKTKTICSSGIMNSPSPIC 123
 Db 263 PALKSPGEGSMSCVQSAEAFQHSQSFSCGEGFELVGPVHVCHTALCVMTAPFPVCKAL 322

RESULT 12

095LG3 PRELIMINARY; PRT; 485 AA.
 AC 095LG3;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE E-SELECTIN.
 OS Odocolleus hemionus (Mule deer) (Black-tailed deer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
 OC Cervidae; Odocolleinae; Odocolleus.
 OX NCBI_TaxID=9872;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21421234; PubMed=11529941;
 RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,
 RA MacLachlan N.J.;
 RT "Characterization of equine E-selectin."
 RL Immunology 103:498-504(2001).
 DR EMBL; AF307970; AK48710.1;-
 KW Lectin; Selectin.
 SQ SEQUENCE 485 AA; 53247 MW; 69959199EAFE9980 CRC64;

Query Match

Best Local Similarity 39.6%; Score 290.5; DB 6; Length 485;
 Matches 53; Conservative 19; Mismatches 48; Indels 1; Gaps 1;

QY 1 CEPLAPELLGTMDCTHHPGNFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 60
 Db 181 CQAKHPEHGHVLC-NPLGKFTYNSSCISCAEGYLPSTEARCKMSSGIMSTPLPCNV 238
 QY 61 IQCEPLAPDGLGIMNCHPLASFSFTSACTFTCSGTELLIGKKTKTICSSGIMNSPSPIC 120
 Db 240 VKDALSPDNGVYVNCSSQNGNSLPMWNTTCFECQEGYKILGPDLQCTSSGIMDNKQPTC 299
 QY 121 Q 121
 Db 300 K 300

RESULT 13

095LG2 PRELIMINARY; PRT; 484 AA.
 AC 095LG2;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE E-SELECTIN.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21421234; PubMed=11529941;
 RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,
 RA MacLachlan N.J.;
 RT "Characterization of equine E-selectin."
 RL Immunology 103:498-504(2001).
 DR EMBL; AF307971; AK48711.1;-
 KW Lectin; Selectin.

SQ SEQUENCE 484 AA; 52951 MW; 617930C1C2F47B44 CRC64;

Query Match

Best Local Similarity 37.6%; Score 275.5; DB 6; Length 484;
 Matches 49; Conservative 21; Mismatches 50; Indels 1; Gaps 1;

QY 1 CEPLAPELLGTMDCTHHPGNFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 60
 Db 180 CQAKHPEHGHVLC-NPLGKFTYNSSCISCAEGYLPSTEARCKMSSGIMSTPLPCNV 238

QY 61 IQCEPLAPDGLGIMNCHPLASFSFTSACTFTCSGTELLIGKKTKTICSSGIMNSPSPIC 120
 Db 239 VKDALSPDNGVYVNCSSQNGNSLPMWNTTCFECQEGYKILGPDLQCTSSGIMDNKQPTC 298

QY 121 Q 121
 Db 299 K 299

RESULT 14

028982 PRELIMINARY; PRT; 482 AA.
 AC 028982;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE E-SELECTIN.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97075911; PubMed=8918234;
 RA Winkler H., Brostjan C., Csizmadia V., Natarajan G., Anrather J.,
 RA Bach F.H.;
 RT "The intron-exon structure of the porcine E-selectin-encoding gene."
 RL Gene 176:67-72(1996).
 DR EMBL; U37521; AAC48680.1;-
 DR HSSP; P16581; IESL.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001304; lectin_c.
 DR InterPro; IPR002396; Selectin.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PR00008; EGF_1.
 DR Pfam; PR00059; lectin_c_1.
 DR Pfam; PR00084; sushi_4.
 DR PRINTS; PR00343; SELECTIN.
 DR SMART; SM00032; CCP_4.
 DR SMART; SM00034; CLECT_1.
 DR SMART; SM00181; EGF_1.
 DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 1.
 KW EGF-like domain; Glycoprotein; Lectin; Selectin.
 SQ SEQUENCE 482 AA; 52341 MW; 97DC5D70B115944 CRC64;

Query Match

Best Local Similarity 33.4%; Score 245; DB 6; Length 482;
 Matches 43; Conservative 22; Mismatches 56; Indels 2; Gaps 2;

QY 1 CEPLAPELLGTMDCTHHPGNFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 59
 Db 240 CDTVGHDPNGDVSNCHSSISGEFAVYKSTCHFTCAEGFGLOGPAOICTAOGQWTOQAPVCK 299
 QY 60 VIQCEPLAPDGLGIMNCHPLASFSFTSACTFTCSGTELLIGKKTKTICSSGIMNSPSPIC 118
 Db 300 AKKCAVSPKNGLVKFTHTSPGTGFTYSSCAFSCEBGEFLKGSQDLACTSOGQWTOQAPVCK 359
 QY 119 IQ 121

DB 360 SCQ 362

RESULT 15

Q9ES77 PRELIMINARY; PRT; 3567 AA.

AC Q9ES77:

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE POLYDOM PROTEIN PRECURSOR.

GN POLYDOM.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C3H/HENSIC;

RX G11ges D., Vinit M.-A., Callebaut I., Coulombel L., Cacheux V.,

RA Romeo P.-H., Vigon I.;

RT "Polydom : a secreted protein with pentraxin, complement control

RT protein, epidermal growth factor and von Willebrand factor A

RT domains.";

RL Blochem. J. 352:49-59(2000).

DR EMBL: AF206329; AAC32160.1; -.

DR HSSP: P00740; 1EDM.

DR MGI: 1928849; Polydom.

DR InterPro: IPR000152; Asx_hydroxyl.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR000742; EGF_2.

DR InterPro: IPR001881; EGF_Ca.

DR InterPro: IPR001438; EGF-II.

DR InterPro: IPR003410; HYR.

DR InterPro: IPR001759; Pentaxin.

DR InterPro: IPR000436; Sushl_SCR_CCP.

DR InterPro: IPR002035; VWFA.

DR Pfam: PF00008; EGF; 10.

DR Pfam: PF02494; HYR; 2.

DR Pfam: PF00084; sushl; 33.

DR Pfam: PF00092; vwa; 1.

DR PRINTS: PR00010; EGFBLD.

DR PRINTS: PR00895; PENTAXIN.

DR PRINTS: PR00453; VWFADOMAIN.

DR PRODOM: PD002153; Pentaxin; 1.

DR SMART: SM00032; CCP; 34.

DR SMART: SM00181; EGF_CA; 9.

DR SMART: SM00179; EGF_Like; 3.

DR SMART: SM00159; PTH; 1.

DR SMART: SM00327; VMA; 1.

DR PROSITE: PS0010; ASX_HYDROXYL; UNKNOWN_1.

DR PROSITE: PS50234; VWFA; 1.

KW Signal.

FT SIGNAL.

SO SEQUENCE 3567 AA; 387391 MW; 8FBA8276E12293E5 CRC64;

Query Match 26.2%; Score 192; DB 11; Length 3567;

Best Local Similarity 27.1%; Pred. No. 1.7e-13;

Matches 49; Conservative 17; Mismatches 51; Indels 64; Gaps 5;

QY 1 CEPLEAPELGTMDCTHPFG-NFSFSSQCAFSCSEGTNLGTIEETTCGPGNMNSPEPTCQ 59

DB 2140 CIPVRCGEPPSIANGYPSGTINSGFAVAVAYSCHKGFYIKGEKSKVCEATGQSKPTPTCH 2199

QY 60 VIQCE-----PLSNP 69

DB 2200 PVSCKNEPPKVENGFLEHTTGRTFESEARQCNPGYKAAGSPVFCOANRHMHDAPLSCOT 2259

QY 70 DLGINCSHPL-----ASFSTSACTFCISGTGLICKKTKICSSGIWS-NPSPI 119

DB 2260 PL---NCKKPPPTQNGFLKGESEFVGSKYQFVYCNQGYELVGNNSWTQCKSKGSKMPPSK 2316

QY 120 C 120

DB 2317 C 2317

RESULT 16

Q917E3 PRELIMINARY; PRT; 866 AA.

AC Q917E3:

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE HIG PROTEIN.

GN HIG OR CG2040.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abtil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,

RA Butts K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heitman J.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,

RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,

RA Palazzolo M., Plattman G.S., Pan S., Pollard J., Part V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Stong R., Sun E.,

RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M.C., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003834; AAC22297.1; -.

DR HSSP: P10998; 1YVD.

DR FLYBase: FBgn0010114; h1g.

DR InterPro: IPR003599; Ig.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR002396; Selectin.

DR InterPro: IPR000436; Sushl_SCR_CCP.

DR Pfam: PF00047; Ig; 1.

DR Pfam: PF00084; sushl; 4.


```

Query Match          25.0%: Score 183; DB 5; Length 958;
Best Local Similarity 30.3%: Pred. No. 4.4e-13;
Matches 46; Conservative 11; Mismatches 33; Indels 42; Gaps 4;

QY      4 LEAPELGTMDCITHPGG-----NFSFSQCAECSCEG 34
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      685 IENNDSGCTCTSPRGCLNSTAVVATGTCPOLTEPLAPLKRLGKNKQGQAHNECPRG 744
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      35 TNLGIEETTCGPEGSNMSSPEPTCOVIOCE--PLSAPDLGIMNCSHPLASFSTSA--C 89
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      745 FRLGAMNATATCLASGNMSSPTPTCHAIQCPRLDDPHILIELN-----TSAMGRA 796
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      90 TFICTSETELGIGKKTKTICSSGIMNSPPIQ 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      797 VFKCQWGFKLITGPALDCEPDSGWSGVPVPRCK 828

RESULT 19
Q9NU87 PRELIMINARY; PRT: 1172 AA.
AC Q9NU87;
DT 01-OCT-2000 (Tremblrel. 15, created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-DEC-2001 (Tremblrel. 15, last annotation update)
DE D1177P10.1.1 (H FACTOR 1 (COMPLEMENT) ISOFORM 1).
GN HFI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA BIRD C.;
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL, AL049744; CAB70597.1; -.
DR HSSP; P08603; 1HPH.
DR InterPro; IPR000436; Sush1_SCR_CGP

```

[illegible]

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Glibos R.A., Myers R.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003663; AAF53815.2; -
 DR HSSP: 046655; ICRH.
 DR FLYBase: FBgn0032797; CG10186.
 DR InterPro: IPR000923; Copper_blue.
 DR InterPro: IPR000566; Lipoclin_cyFABP.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR002221; WAP.
 DR Pfam: PF00084; sushi; 14.
 DR Pfam: PF00095; wap; 1.
 DR SMART: SM00032; CCP; 17.
 DR SMART: SM00217; WAP; 1.
 DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
 DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
 SQ SEQUENCE 1633 AA; 180864 MW; 174E7775C51FD60F CRC64;

Query Match 23.6%; Score 173; DB 5; Length 1653;
 Best Local Similarity 25.4%; Pred. No. 1.3e-11;
 Matches 48; Conservative 16; Mismatches 49; Indels 76; Gaps 5;

OY 1 CEPLAELGTMDCTHPFGNFSFSS-----QCAFSCSEGTNLGIEETTCGP 47
 DB 1437 CESVEECGDI-----PLGMSNASPRVSVLSREVGRAAFSCAGYGLGPAEALICNP 1489
 OY 48 FGNSSPEPTCOVIOCEPLSAPDGL----- 72
 DB 1490 TGEWSAPLPCTVEQCPNGAPQNGVAGSAPYAGDVQFNPCEYMGQPIIACODN 1549
 OY 73 -----IMNCSHP-----LASFSEPTSA---CTFICSEGTLLGKKTICSSG 111
 DB 1550 ARMSGGLPKCVQACSYRGTYVISGRMSSVKFYFALGSEITTFCDAGLDLRGSKVLACLKNG 1609
 OY 112 TWSNPSPTC 120
 DB 1610 KWSAIPTC 1618

RESULT 21
 O9H284 PRELIMINARY; PRT; 481 AA.
 AC O9H284;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE SEROLOGICALLY DEFINED BREAST CANCER ANTIGEN NY-BR-38.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BREAST;
 RA Seanlan M.J., Gout I., Stockert E., Gure A.O., Jaeger D., Chen Y.-T.,
 RA Old L.J.;
 RT "Humoral Immunity to Human Breast Cancer: Antigen Definition and
 RT Quantitative Analysis of mRNA Expression.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF308289; AAG48257.1; -
 DR HSSP: P10998; IYVD.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR Pfam: PF00084; sushi; 7.
 DR SMART: SM00032; CCP; 8.
 SQ SEQUENCE 481 AA; 51891 MW; FD99724DEF7C69B0B CRC64;

Query Match 23.5%; Score 172.5; DB 4; Length 481;
 Best Local Similarity 31.2%; Pred. No. 3.6e-12;
 Matches 39; Conservative 18; Mismatches 59; Indels 9; Gaps 4;

OY 1 CEPLAELGTMDCTHPFGNFSFSSQCAFSCSEGTNLGIEETTCGPFGNSSPEPTCOV 60
 DB 107 CTP---PPLISFVGPRISSALHFGSTYKYSVGSGFFLRGNSTLTLCQDGDGMSSPLPCVP 163
 OY 61 IQC-EPLSAPDLGIMNCSHPPLASFSTACTFTICSEGTLLGKKTICSSGIMNSDPT 119
 DB 164 VECQPEIPIFN-GIID----VOGLAYLSTALYTCRPGFELVGNNTTTLGCHGHWLGCKPT 218
 OY 120 COKLD 124
 DB 219 CKATE 223

RESULT 22
 O91YB6 PRELIMINARY; PRT; 1236 AA.
 AC O91YB6;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE COMPLEMENT INHIBITORY FACTOR H.
 GN FH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-LIVER;
 RA Demberg T., Goetze O., Schlar G.;
 RT "Rat complement factor H: molecular cloning, sequencing and expression
 RT in tissues and isolated cells";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ320522; CAC67513.1; -
 SQ SEQUENCE 1236 AA; 140343 MW; 1AC89FPA28232EBF CRC64;

Query Match 23.1%; Score 169; DB 11; Length 1236;
 Best Local Similarity 28.8%; Pred. No. 2.8e-11;
 Matches 36; Conservative 26; Mismatches 47; Indels 16; Gaps 5;

OY 14 CTH----PFGN-----FSSQCAFSCSEGTNLG-IEETTCGPFGNSSPEPTCOV 60
 DB 85 CGHGDPTFGSFLRAWSEFEFGAKVYVTCDEGYQLGEIDYRECDADG-WINDIPICEV 143
 OY 61 IOCEPLSAPDLG--IMNCSHPPLASFSTACTFTICSEGTLLGKKTICSSGIMNSDPT 118
 DB 144 VKCLPVELENGRIVSGAEPDDEYFFGOVYRECNCSGFKIEGOKHEHCSENGLSNKP 203
 OY 119 ICOKL 123
 DB 204 QCYEI 208

RESULT 23
 O96JUT7 PRELIMINARY; PRT; 1316 AA.
 AC O96JUT7;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE CDNA FLJ14964 FIS, CLONE PLACE4000581, MODERATELY SIMILAR TO
 DE FIBROPELIN I PRECURSOR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PLACENTA;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagehara K., Masuko Y., Sasaki N.,
 RT "NEBO human cDNA sequencing project";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK027870; BAB55420.1;
 SQ SEQUENCE 1316 AA; 144524 MW; 80615BBA3A4F00A5 CRC64;

Query Match 23.0%; Score 168.5; DB 4; Length 1316;
 Best Local Similarity 30.9%; Pred.No. 3,4e-11;
 Matches 42; Conservative 15; Mismatches 68; Indels 11; Gaps 4;

OY 1 CEPLAPELGTW---DCTHPGNEFSFSQCAFSCESEGTNLGTGIEETTCGPGNWSSEPT 57
 DB 145 CALKPPENGFIQNTC-----NNHFNACGYRCHGPFVLVSSITLCIPNGLMSGLESY 199
 OY 58 CQVIOCEPLAPDLGIMNCSHPLASFSTSACTFICSECTELIGKRTICSSGIMSNPS 117
 DB 200 CWRRTCPHLRQPKHGHISCS--TREMLYKTTCLVACDEGYLEGSDKLTCQGNQWDGPE 257
 OY 118 PIC-QKLKSPSMIKE 132
 DB 258 PRCVERHCSTFQMPKD 273

RESULT 24

OY 029530 PRELIMINARY: PRT: 2014 AA.
 AC 029530:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE COMPLEMENT RECEPTOR 1 (FRAGMENT).
 GN CRI.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94292799; PubMed=8021505;
 RA Birmingham D.J., Shen X.P., Hourcade D., Nickells M.W., Atkinson J.P.;
 RT "Primary sequence of an alternatively spliced form of CRI. Candidate
 RT for the 75,000 M(r) complement receptor expressed on chimpanzee
 RT erythrocytes";
 RL J. Immunol. 153:691-700(1994).
 DR EMBL: L24920; AAS1438.1;
 DR HSSP: P08603; 1HFI.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR InterPro: IPR000834; Zn_cathopept.
 DR Pfam: PF00084; Sush1; 30.
 DR SMART: SM00032; CCP; 30.
 DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
 DR PROSITE: PS00087; SOD_CU_ZN_1; UNKNOWN_1.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 1
 SO SEQUENCE 2014 AA; 221281 MW; 6D6C3A74D81F1DB9 CRC64;

Query Match 22.9%; Score 167.5; DB 6; Length 2014;
 Best Local Similarity 32.6%; Pred.No. 7,4e-11;
 Matches 46; Conservative 16; Mismatches 56; Indels 23; Gaps 6;

OY 1 CEPLAPELGTW---DCTHPGNEFSFSQCAFSCESEGTNLGTGIEETTCGPGN--WSSPEPT 57
 DB 1432 CQPPPEPFGMW---HINDTDFGSTVNVSCNCGFRLLGSPPTTCLVSCNNVTWKKAPI 1488
 OY 58 CQVIOCEP---ISAPDLGIMNCSHPLASFSTSACTFICSEST-----ELIGKRTICE 108
 DB 1489 CEIICEPPTISNGDF---YSNNRASFRHNGTAVTYOCHTGPDEQELFELVGERSIYCT 1544

OY 109 SS---GIWNSPITCOKLIDK 125
 DB 1545 SKDDQGVWSSPPRCISITNK 1565

RESULT 25
 OY 016745 PRELIMINARY: PRT: 2039 AA.
 AC 016745:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE COMPLEMENT RECEPTOR 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94065175; PubMed=8245463;
 RA Vik D.P., Wong W.W.;
 RT "Structure of the gene for the F allele of complement receptor type 1
 RT and sequence of the coding region unique to the S allele.";
 RL J. Immunol. 151:6214-6224(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Vik D.P., Wong W.W.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L17418; AAB60694.1;
 DR EMBL: L17390; AAB60694.1; JOINED.
 DR EMBL: L17399; AAB60694.1; JOINED.
 DR EMBL: L17409; AAB60694.1; JOINED.
 DR EMBL: L17419; AAB60694.1; JOINED.
 DR EMBL: L17420; AAB60694.1; JOINED.
 DR EMBL: L17421; AAB60694.1; JOINED.
 DR EMBL: L17422; AAB60694.1; JOINED.
 DR EMBL: L17423; AAB60694.1; JOINED.
 DR EMBL: L17391; AAB60694.1; JOINED.
 DR EMBL: L17392; AAB60694.1; JOINED.
 DR EMBL: L17393; AAB60694.1; JOINED.
 DR EMBL: L17394; AAB60694.1; JOINED.
 DR EMBL: L17395; AAB60694.1; JOINED.
 DR EMBL: L17396; AAB60694.1; JOINED.
 DR EMBL: L17397; AAB60694.1; JOINED.
 DR EMBL: L17398; AAB60694.1; JOINED.
 DR EMBL: L17400; AAB60694.1; JOINED.
 DR EMBL: L17401; AAB60694.1; JOINED.
 DR EMBL: L17402; AAB60694.1; JOINED.
 DR EMBL: L17403; AAB60694.1; JOINED.
 DR EMBL: L17404; AAB60694.1; JOINED.
 DR EMBL: L17405; AAB60694.1; JOINED.
 DR EMBL: L17406; AAB60694.1; JOINED.
 DR EMBL: L17407; AAB60694.1; JOINED.
 DR EMBL: L17408; AAB60694.1; JOINED.
 DR EMBL: L17410; AAB60694.1; JOINED.
 DR EMBL: L17411; AAB60694.1; JOINED.
 DR EMBL: L17412; AAB60694.1; JOINED.
 DR EMBL: L17413; AAB60694.1; JOINED.
 DR EMBL: L17414; AAB60694.1; JOINED.
 DR EMBL: L17415; AAB60694.1; JOINED.
 DR EMBL: L17416; AAB60694.1; JOINED.
 DR EMBL: L17417; AAB60694.1; JOINED.
 DR HSSP: P08603; 1HFI.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR InterPro: IPR000834; Zn_cathopept.
 DR Pfam: PF00084; Sush1; 30.
 DR SMART: SM00032; CCP; 30.
 DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
 DR PROSITE: PS00087; SOD_CU_ZN_1; UNKNOWN_1.
 KW Receptor.
 SO SEQUENCE 2039 AA; 223603 MW; B82FCB11CG816635 CRC64;

[illegible]

OY 101 GKKTICSSGIMNSPSPIC 120
 DB 2779 GSSALTCMANGIMDRSLPKC 2798

RESULT 28

O96RM4 PRELIMINARY: PRT: 3508 AA.
 AC 096RM4:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CUB AND SUSHI MULTIPLE DOMAINS 1 PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21365705; PubMed=11472063;
 RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,
 RA Sunwoo J.B., Gollin S.M., Scholnick S.B.:
 RA "Transcript map of the 8p23 putative tumor suppressor region.";
 RL Genomics 75:17-25(2001).
 DR EMBL: AF333704; AAK73475.1;
 DR SEQUENCE 3508 AA; 382824 MW; 9268C3B8F3F78C18 CRC64;

Query Match 21.8%; Score 160; DB 4; Length 3508;
 Best Local Similarity 26.4%; Pred. No. 1.1e-09;
 Matches 37; Conservative 17; Mismatches 48; Indels 38; Gaps 4;

OY 5 EAPELGTMDCTHPFGN-----FSFSQAFSCSCTNLGTIEETTCGPFQWSSP 54
 DB 2673 QPVCVPTICGHP-GNPAHGTNGSEFNLNDVNFCTNGYLQGSRAQCSNSQWSSP 2731
 OY 55 EPTCOVIOCEPLSAPDLGIMNCSHP-----LASFSFTSACTFICSEGTETI 100
 DB 2732 LPTCKVY-----NCSDFQFVENAIRHQQNPESFEIGMSILYHCKKGFHL 2778
 OY 101 GKKTICSSGIMNSPSPIC 120
 DB 2779 GSSALTCMANGIMDRSLPKC 2798

RESULT 29

O921P0 PRELIMINARY: PRT: 390 AA.
 AC 0921P0:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SIMILAR TO DECAV ACCELERATING FACTOR 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.:
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC011314; AAH11314.1;
 DR SEQUENCE 390 AA; 42637 MW; 3FB68595F07A67FF CRC64;

Query Match 21.4%; Score 156.5; DB 11; Length 390;
 Best Local Similarity 30.9%; Pred. No. 2.3e-10;
 Matches 42; Conservative 16; Mismatches 63; Indels 15; Gaps 6;

OY 1 CEPLEAPELGTMDCTH---PFGNFSFSQAFSCSEGTNLGTIEETTCGPFQWSSP 54
 DB 158 CRRKSCPFRKIDNGHINIPFG-ILFGSEINFSQNGRYLGVSVFSCSVTGNTVDMDE 216

OY 55 EPTCOVIOCEPLSAPDLGIMNCSHPPLASFPSTACTFICSEGTILGKKTICSS- 110
 DB 217 FVCTETIHCPPEPKKINNGIMRGESD--SYTISQVYVTSQDKGVLYGNASTYCTVSKSOV 274

OY 111 GIMNSPSPICQKLDKS 126
 DB 275 GOWSSPPPRC--IEKS 288

RESULT 30

O08569 PRELIMINARY: PRT: 533 AA.
 AC 008569:
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ACROSOMAL MATRIX COMPONENT AM67 PRECURSOR.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HARTLEY; TISSUE=TESTIS;
 RX MEDLINE=97284752; PubMed=9139729;
 RA Foster J.A., Friday B.B., Maulit M.T., Blobel C., Wintrey V.P.,
 RA Olson G.E., Kim K.S., Gerton G.L.:
 RA "AM67, a secretory component of the guinea pig sperm acrosomal matrix,
 RT is related to mouse sperm protein sp56 and the complement component 4-
 RT binding proteins.";
 RL J. Biol. Chem. 272:12714-12722(1997).
 DR EMBL: U75654; AAC13888.1;
 DR HSP; P10998; IYVD.
 DR InterPro: IPR00436; Sush1_SCR_CCP.
 DR Pfam: PF00084; Sush1; 7.
 DR SMART: SM00032; CCP; 7.
 DR KX SIGNAL.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 533 ACROSOMAL MATRIX COMPONENT AM67.
 SO SEQUENCE 533 AA; 59772 MW; EDBDD487A45389 CRC64;

Query Match 21.1%; Score 155; DB 11; Length 533;
 Best Local Similarity 31.5%; Pred. No. 4.9e-10;
 Matches 35; Conservative 20; Mismatches 42; Indels 14; Gaps 5;

OY 23 FSSQAFSCSEGTNLGTIEETTC--GPGNMSPPPTGVOIOCEPLSAPDLGIMNCSHP 79
 DB 110 FGTIEFSCSKGSLIGSTTSCQSGKTVDMSDPLPECVIYKCD--SPPD--ISNGKHS 165
 OY 80 LAS---FSTTACTFICSEGTILGKKTIC---ESSGIMNSPSPICQK 123
 DB 166 GDEUDELTYGLVTVCDPNYSILGNASISCLVANKTVGWSNPTCKV 216

RESULT 31

O99JAI PRELIMINARY: PRT: 555 AA.
 AC 099JAI:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE C4BP ALPHA-CHAIN PRECURSOR.
 GN C4BPA.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EPIDIDYMIS. AND LIVER;
 RX MEDLINE=21154058; PubMed=11254714;
 RA Nonaka M.I., Wang G., Mori T., Okada H., Nonaka M.;

```
RT "Novel androgen-dependent promoters direct expression of the C4b-
RT binding protein alpha-chain gene in epididymis."
RL J. Immunol. 166:4570-4577(2001).
DR EMBL: AB049465; BAB39737.1; -.
DR EMBL: AB049466; BAB39738.1; -.
DR EMBL: AB049467; BAB39739.1; -.
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 8.
DR SMART: SM00032; CCP; 8.
KM Signal.
FT SIGNAL 1 13 POTENTIAL.
FT CHAIN 14 555 POTENTIAL.
SQ SEQUENCE 555 AA; 61628 MW; 7AFA5462AFC7B6AC CRC64;

Query Match 21.1%; Score 155; DB 11; Length 555;
Best Local Similarity 33.9%; Pred. No. 5.1e-10;
Matches 37; Conservative 18; Mismatches 40; Indels 14; Gaps 5;

OY 25 SOCAFSCSEGTNLGIEETTCGPGF---NMSSPEPTCVOICPELSPADLGINCSHPLA 81
ID 019122 PRELIMINARY; PRT; 222 AA.
AC 019122;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE MEMBRANE COFACTOR PROTEIN (FRAGMENT).
GN CD46.
OS Callimico goeldii (Goeldi's marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callimico.
OX NCBL_TaxID=9495;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=97366679; PubMed=9223509;
RA Hsu E.C., Dorig R.E., Sarangl F., Marcell A., Iorio C.,
RA Richardson C.D.;
RT "Artificial mutations and natural variations in the CD4 molecules
RT from human and monkey cells define regions important for measles virus
RT binding."
RL J. Virol. 71:6144-6154(1997).
DR EMBL: U87916; AAB66816.1; -.
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 3.
DR SMART: SM00032; CCP; 3.
FT NON_TER 222 222
SQ SEQUENCE 222 AA; 24523 MW; 7CD336D0620CC0E9 CRC64;

Query Match 21.1%; Score 154.5; DB 6; Length 222;
Best Local Similarity 31.5%; Pred. No. 2e-10;
Matches 39; Conservative 19; Mismatches 51; Indels 15; Gaps 5;

OY 19 GNFPSGQAFSCSEGTNLGIEETTC---GPGNWSPEPTCVOICPELSPADLGIN 75
ID 029528 PRELIMINARY; PRT; 1911 AA.
AC 029528;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE COMPLEMENT COMPONENT RECEPTOR TYPE 1 (FRAGMENT).
GN CRI.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciae;
OC Cercopitheciae; Papio.
OX NCBL_TaxID=9571;
RN 11
RP SEQUENCE FROM N.A.
RA Clemenza L., Subramanian B.V., Nickells M.W., Hourcade D.E.,
RA Atkinson J.P.;
RT "Primary sequence of the baboon 200 kDa C3b/C4b receptor (CRI).";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L39791; AAA62170.1; -.
DR HSSP: P08603; 1HFI.
DR InterPro: IPR001424; SOD_CU_ZN.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR000834; Zn_carboxypept.
DR Pfam: PF00084; sushi; 29.
DR SMART: SM00032; CCP; 29.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_4.
DR PROSITE: PS00087; SOD_CU_ZN_1; UNKNOWN_1.
KM Receptor.
FT NON_TER 1911 1911
SQ SEQUENCE 1911 AA; 210174 MW; 535AADD0EAF521D CRC64;

Query Match 20.9%; Score 153; DB 6; Length 1911;
Best Local Similarity 32.8%; Pred. No. 3.7e-09;
Matches 39; Conservative 14; Mismatches 46; Indels 20; Gaps 5;

OY 23 FSSQAFSCSEGTNLGIEETTCGPGN---WSSPEPTCVOICPE---LSAPDLGIN 76
ID 092313 PRELIMINARY; PRT; 3564 AA.
AC 092313;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CSMO1.
GN CSMO1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBL_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRATN-C57BL/6;
RX MEDLINE=21365705; PubMed=11472063;
RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,
RA Sunwoo J.B., Gollin S.M., Schoinick S.B.;
RT "Transcript map of the bp23 putative tumor suppressor region."
RL Genomics 75:17-23(2001).
DR EMBL: AV017475; AAG54083.1; -.

OY 128 SMIK 131
```

```
Db 168 PVIE 171

RESULT 33
OY 23 FSSQAFSCSEGTNLGIEETTCGPGN---WSSPEPTCVOICPE---LSAPDLGIN 76
ID 092313 PRELIMINARY; PRT; 3564 AA.
AC 092313;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CSMO1.
GN CSMO1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBL_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRATN-C57BL/6;
RX MEDLINE=21365705; PubMed=11472063;
RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,
RA Sunwoo J.B., Gollin S.M., Schoinick S.B.;
RT "Transcript map of the bp23 putative tumor suppressor region."
RL Genomics 75:17-23(2001).
DR EMBL: AV017475; AAG54083.1; -.

OY 77 SHPLASFTSACTFTCSCT-----ELIGKKTICSS-----GIMSNSPTCKLDK 125
ID 1419 SNMRTSFHSGTVVYTOCHGPDGDLFELVGERSTYCTSKDOVGAMSSPPRCISTNK 1477
DB 1419 SNMRTSFHSGTVVYTOCHGPDGDLFELVGERSTYCTSKDOVGAMSSPPRCISTNK 1477

RESULT 34
OY 23 FSSQAFSCSEGTNLGIEETTCGPGN---WSSPEPTCVOICPE---LSAPDLGIN 76
ID 092313 PRELIMINARY; PRT; 3564 AA.
AC 092313;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CSMO1.
GN CSMO1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBL_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRATN-C57BL/6;
RX MEDLINE=21365705; PubMed=11472063;
RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,
RA Sunwoo J.B., Gollin S.M., Schoinick S.B.;
RT "Transcript map of the bp23 putative tumor suppressor region."
RL Genomics 75:17-23(2001).
DR EMBL: AV017475; AAG54083.1; -.

OY 128 SMIK 131
```

SQL SEQUENCE 3564 AA; 387865 MW; 70824C55B0674609 CRC64;

Query Match 20.9%; Score 153; DB 11; Length 3564;
Best Local Similarity 26.7%; Pred. No. 7.7e-09;
Matches 35; Conservative 17; Mismatches 59; Indels 20; Gaps 4;

QY 1 CEPLAEPGLGMDCTHPFNG-----FSFSQCAFSCSGTNLTGIEETTCGPGFN 50
DB 2733 CVP1-----TCGHP-GNPAHGLTNGTEFNLNDLVNFTCHTGRLOQASRAQCSNGQ 2783
QY 51 WSSPEPTCOVICPEPLSAPDLGIMNSHPL-ASFSTSACTFICSEGTGELIGKKTICSS 109
DB 2784 WSSPLPICVNCSDPGSVENAVRHQGNPPESEFVGTGVTWYHCKTGIFVLLSSALTCA 2843
QY 110 SG1WSNPSPIC 120
DB 2844 SG1MDRSLPKC 2854

RESULT 35

ID 002839 PRELIMINARY; PRT; 363 AA.

AC 002839; PRELIMINARY; PRT; 363 AA.
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PORCINE MEMBRANE COFACTOR PROTEIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxId=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97343414; PubMed=9199970;
RA Toyomura K., Fujimura T., Murakami H., Natsume T., Shigenisa T.,
RA Inoue N., Takeda J., Kinoshita T.;
RT "Molecular cloning of a pig homologue of membrane cofactor protein
RT (CD46).";
RL Int. Immunol. 9:869-876(1997).
DR EMBL; D70897; BAA20476.1; -.
DR HSSP; P10998; 1VD.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; sush1. 4.
DR SMART: SM00032; CCP; 4.
SQ SEQUENCE 363 AA; 39692 MW; 8A5FF329B8E03153 CRC64;

Query Match 20.8%; Score 152.5; DB 6; Length 363;
Best Local Similarity 29.5%; Pred. No. 6.2e-10;
Matches 38; Conservative 16; Mismatches 64; Indels 11; Gaps 4;

QY 1 CEPLAEPGLGMDCTHPFNGFSFSQCAFSCSGTNLTGIEETTCGPGFN--WSSPEPT 57
DB 108 CSNLPDLNGQV--SYNCDMLFGSKAQFTCMNGFIYAETVYQCVSNVAMWSPSP 165
QY 58 CQVIOCEPLSAPDLGIMNSHPLSFTSACTFICSEGT-----ELIGKKTICSSGI 112
DB 166 CEKILCKPGEIPIGKRYTNSHK-DVEFYENVYYSCLSTGDEFSLVGESSLFCIGKDE 224
QY 113 WSNPSPIC 121
DB 225 WSDPPECK 233

RESULT 36

ID 09J2M6 PRELIMINARY; PRT; 395 AA.

AC 09J2M6; PRELIMINARY; PRT; 395 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE COMPLEMENT BINDING PROTEIN.

OS Macaca mulatta rhadinovirus 26-95.
OC Gammae: dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxId=119193;

QY 21 FSFSQCAFSCSGTNLTGIEETTC-----GPGMNSPEPTCOVIOCEPLSAPDLGINN 75
DB 109 FRKSNITKYKNCNTGYLLGATVTRCLKYDSNLVDWQPAAPTCIEKCR--KQPD--IEN 164
QY 76 CS-HPLASF-SFTSACTFICSEGTGELIGKKTICSSG1WSNPSPICXL 123
DB 165 GKXYPVQEFYVNLFTTCNKDPSLIGNTTTCMTGTWSSPVPKCOOI 214
RESULT 37
ID 090611 PRELIMINARY; PRT; 354 AA.
AC 090611;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FURROWED (FRAGMENT).
GN FW.
OS Drosophila ananassae (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7217;
RN [1]
RP SEQUENCE FROM N.A.
RA Marsh B.J., Chen Y., Stephan W.;
RT "Interaction of Natural Selection and Gene Flow in Drosophila
RT ananassae.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF185289; AAF03885.1; -.
DR Flybase; FBgn0014574; Dana.fw.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; sush1. 6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 6.
FT NON_TER 1
FT 354
SQ SEQUENCE 354 AA; 38526 MW; AD02699F25DF9CAB CRC64;

Query Match 20.7%; Score 151.5; DB 5; Length 354;
Best Local Similarity 30.0%; Pred. No. 8e-10;
Matches 39; Conservative 19; Mismatches 57; Indels 15; Gaps 4;

```

OY 5 EAPBLGTMDCTHP-----FG-NFSSSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCOYIOCEPL 55
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 EAPVCELVTCEMPVPVPGSYVYDYDNVMSKIKYSCDDPHIMHGISDLCELDGSGKSDTA 210
OY 56 PTCOYIOCEPLSAPDLGIMNCNHPLASFS--FTSACTFICSGTELIGKKTKICSSSGIWM 113
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 PCEIYIDCGPI---LPIPYGGHKYVNTSYVGSSEVSCIOSHMLSGVKKQCLDSGVW 266
OY 114 SNMSPICOKL 123
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 267 SDASPKCEEI 276

RESULT 38
OY 09HAR2 PRELIMINARY; PRT; 410 AA.
ID 09HAR2:
AC 09HAR2:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE BA524D16A.1 (SUSHI-REPEAT-CONTAINING PROTEIN) (FRAGMENT).
GN SRFX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilson S.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL391688; CAC16060.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR InterPro; IPR003410; HXR.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF02494; HXR; 1.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 410 AA; 46781 MW; D32228E1A1AFAF370 CRC64;

Query Match 20.7%; Score 151.5; DB 4; Length 410;
Best Local Similarity 29.7%; Pred. No. 9.4e-10;
Matches 35; Conservative 13; Mismatches 63; Indels 7; Gaps 3;

OY 10 GTMDCTHPFG---NFSSSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCOYIOCEPL 66
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 12 GEATCYSPKGNVHSSLGTRCELSCDRGFRLIGRRSVQCLPSRWSC-TAYCROMRCHAL 70
OY 67 SAPDLGIMNCNHPLASFSFTSACTFICSEGTTELIGKKTKICSSSGIWMSPICOKLD 124
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 PFITSGYTCTNGVL---LDSRCDYSCSSGYHLEGDRSRICMEDGRMSGGEPVCVDID 125

RESULT 39
OY 060687 PRELIMINARY; PRT; 465 AA.
ID 060687:
AC 060687:
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE SUSHI-REPEAT PROTEIN.
GN SRPUL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuregawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Shinjo T.,
RA Rakestraw K.M., Neave C.W., Look T.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AF060567; AAC15765.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR InterPro; IPR003410; HXR.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF02494; HXR; 1.
DR Pfam; PF00084; sush1; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFB8 CRC64;

Query Match 20.7%; Score 151.5; DB 4; Length 465;
Best Local Similarity 29.7%; Pred. No. 1.1e-09;
Matches 35; Conservative 13; Mismatches 63; Indels 7; Gaps 3;

OY 10 GTMDCTHPFG---NFSSSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCOYIOCEPL 66
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 GEATCYSPKGNVHSSLGTRCELSCDRGFRLIGRRSVQCLPSRWSC-TAYCROMRCHAL 125
OY 67 SAPDLGIMNCNHPLASFSFTSACTFICSEGTTELIGKKTKICSSSGIWMSPICOKLD 124
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 PFITSGYTCTNGVL---LDSRCDYSCSSGYHLEGDRSRICMEDGRMSGGEPVCVDID 180

RESULT 40
OY 022328 PRELIMINARY; PRT; 560 AA.
ID 022328:
AC 022328:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HYPONETICAL 61.6 KDA PROTEIN.
GN T07H6.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequencing of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2016(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Geisel C.;
RL "The sequence of C. elegans cosmid T07H6.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL; U53344; AAA96225.1; -.
DR HSSP; P10998; 1VVD.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 7.
DR SMART; SM00032; CCP; 3.
KW Hypothetical protein.
SQ SEQUENCE 560 AA; 61619 MW; FB8923BAC1B320C9 CRC64;

Query Match 20.5%; Score 150.5; DB 5; Length 560;
Best Local Similarity 26.6%; Pred. No. 1.8e-09;
Matches 33; Conservative 24; Mismatches 62; Indels 5; Gaps 2;

OY 1 CEPLPELGTMDCTHPFGN-FSFSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCO 59
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 CKRACPDGDIENGRLREDTFEYPHHVKYSNCPGLTGWSTSRCCSSNGEWTNPANCK 185

```


OY	60	VIOCEPLSAPDLGIMNC	SHPLASF	SFTSACTFIC	SEGT	ELIGKKK	TTIC	ESSG	IMSN	PSPI	119
Db	186	ATECSRPSPLHGK	VGS----	SLTYQSV	TYSCDH	GYRLV	GOVOR	ICIA	ESIM	WGNEPR	241
OY	120	CQKL	123								
Db	242	CEEI	245								

Search completed: September 7, 2002, 10:23:24
 Job time: 552 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:43 ; Search time 18.65 Seconds

(without alignments)
274.047 Million cell updates/sec

Title: US-09-119-209-2_COPY_197_328

Perfect score: 733

Sequence: 1 CEPLAPELLCTMDCTHPFGN.....WSNPSPICQKDKFSMIKE 132

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	727	99.2	372	1	LEM1_HUMAN
2	727	99.2	372	1	LEM1_PANTR
3	724	98.8	372	1	LEM1_PONPY
4	690	94.1	372	1	LEM1_MACMU
5	690	94.1	372	1	LEM1_PAPHA
6	537	73.3	370	1	LEM1_BOVIN
7	535	73.0	372	1	LEM1_RAT
8	518	70.7	372	1	LEM1_MOUSE
9	315	43.0	830	1	LEM3_HUMAN
10	313	42.7	768	1	LEM3_RAT
11	305.5	41.7	612	1	LEM2_MOUSE
12	300	40.9	611	1	LEM2_CANFA
13	300	40.9	768	1	LEM3_MOUSE
14	300	40.9	769	1	LEM3_SHEEP
15	299.5	40.9	549	1	LEM2_RAT
16	296	40.4	610	1	LEM2_HUMAN
17	290	39.6	646	1	LEM3_BOVIN
18	288	39.3	551	1	LEM2_RABIT
19	271.5	37.0	485	1	LEM2_BOVIN
20	245	33.4	484	1	LEM2_PIG
21	178	24.3	958	1	HIG_DROME
22	173	23.6	1231	1	CFAH_HUMAN
23	163.5	22.3	2039	1	CRI_HUMAN
24	160	21.8	1234	1	CFAH_MOUSE
25	155.5	21.2	390	1	DAFL_MOUSE
26	155.5	21.2	1019	1	LFC_TACTR
27	150.5	20.5	1025	1	CR2_MOUSE
28	149.5	20.4	597	1	CARP_MOUSE
29	147.5	20.1	685	1	CFAH_BOVIN
30	146	19.9	518	1	C4BP_RAT
31	145	19.8	610	1	C4BP_BOVIN
32	144.5	19.7	469	1	C4BP_MOUSE
33	144	19.6	258	1	C4BP_RAT

34	142	19.4	507	1	DAF_CAVPO	Q60401 cavia porce
35	141.5	19.3	345	1	APDH_BOVIN	P17690 bos taurus
36	141	19.2	1033	1	CR2_HUMAN	P20023 homo sapien
37	140.5	19.2	764	1	CFAB_HUMAN	P00751 homo sapien
38	139	19.0	263	1	VCP_VACCV	P10998 vaccinia vi
39	139	19.0	381	1	DAF_HUMAN	P08174 homo sapien
40	137	18.7	340	1	DAF_PONPY	P49457 pongo pygma
41	136.5	18.6	377	1	MCP_HUMAN	P15529 homo sapien
42	135.5	18.5	1019	1	LFC_CARRO	Q26422 carcinoscor
43	134	18.3	1627	1	PAPA_HUMAN	Q13219 homo sapien
44	133.5	18.2	407	1	DAF2_MOUSE	Q61476 mus musculu
45	132.5	18.1	752	1	CO2_HUMAN	P06681 homo sapien

ALIGNMENTS

RESULT ID	1	LEM1_HUMAN	STANDARD:	PRT:	372 AA.
AC	1	LEM1_HUMAN	P15023:		
DT	01-JAN-1990	(Rel. 13, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte surface antigen Leu-8) (TQ1) (gp90-MEL)				
DE	(Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).				
GN	SELL OR LYAM1 OR LNH1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89315837; PubMed=2664786;				
RA	Siegelman M.H., Weissman I.L.;				
RT	"Human homologue of mouse lymph node homing receptor: evolutionary conservation at tandem cell interaction domains.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:5562-5566(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Tonsil;				
RX	MEDLINE=89310350; PubMed=2473156;				
RA	Cammerlin D., James S.P., Stamenkovic I., Seed B.;				
RT	"Leu-8/TQ1 is the human equivalent of the MEL-14 lymph node homing receptor.";				
RL	Nature 342:78-82(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89308881; PubMed=2663882;				
RA	Bowen B.R., Nguyen T., Lasky L.A.;				
RT	"Characterization of a human homologue of the murine peripheral lymph node homing receptor.";				
RL	J. Cell Biol. 109:421-427(1989).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90243637; PubMed=1692315;				
RA	Ord D.C., Ernst T.J., Zhou L.J., Rambaldi A., Spertini O.,				
RT	Griffin J., Tedder T.F.;				
RL	"Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils.";				
RN	J. Biol. Chem. 265:7760-7767(1990).				
RP	[6]				
RL	3D-STRUCTURE MODELING.				

```

RX MEDLINE=96074584; PubMed=7488174;
RA Rajorath J., Aruffo A.;
RT "A template for generation and comparison of three-dimensional
RT selectin models.";
RL Biochem. Biophys. Res. Commun. 216:1018-1023(1995).
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENUES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -!- DATABASE: NAME-PROW: NOTE=CD guide CD62L entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd62l.htm"
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M25280; AAC63053.1; -
DR EMBL; X16150; CAA34275.1; -
DR EMBL; X17519; CAA34353.1; -
DR EMBL; X17519; CAA34353.1; ALT_SEQ.
DR EMBL; X16070; CAA34203.1; ALT_SEQ.
DR EMBL; M32414; AAB60700.1; -
DR EMBL; M32406; AAB60700.1; JOINED.
DR EMBL; M32407; AAB60700.1; JOINED.
DR EMBL; M32408; AAB60700.1; JOINED.
DR EMBL; M32409; AAB60700.1; JOINED.
DR EMBL; M32410; AAB60700.1; JOINED.
DR EMBL; M32411; AAB60700.1; JOINED.
DR EMBL; M32412; AAB60700.1; JOINED.
DR EMBL; M32413; AAB60700.1; JOINED.
DR PIR; A33912; A33912.
DR PIR; A34015; A34015.
DR PIR; S06798; S06798.
DR PIR; J10104; J10104.
DR PDB; 1KJB; 03-APR-96.
DR GlycoSuiteDB; P14151; -
DR MIM; 153340; -
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sushin_SCR_CCP.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00059; lectin_c. 1.
DR Pfam; PF00084; sushin_2.
DR PRINTS; PRO0343; SELECTIN.
DR SMART; SM00032; CCP_2.
DR SMART; SM00034; CLECT_1.
DR SMART; SM00181; EGF_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS00041; C-TYPE LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
FT Selection; Signal; Sushin; Repeat; 3d-structure.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155

```

```

FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 37 37 D -> Y (IN REF. 4).
FT CONFLICT 178 178 Y -> H (IN REF. 4).
FT CONFLICT 193 193 F -> L (IN REF. 1 AND 4).
FT CONFLICT 213 213 P -> S (IN REF. 3).
FT CONFLICT 214 214 L -> F (IN REF. 4).
FT CONFLICT 218 220 SFS -> NEN (IN REF. 2).
FT CONFLICT 242 242 G -> E (IN REF. 2).
SQ SEQUENCE 372 AA; 42187 MW; 6EA918BCA2D3643 CRC64;

Query Match 99.2%; Score 727; DB 1; Length 372;
Best Local Similarity 99.2%; Pred. No. 7,5e-64;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAEPLGTMDCTHPFGNFSSQCAFSCSEGTNLGTIEFTTGGPFGMWSPEPTCOV 60
DB 197 CEPLAEPLGTMDCTHPLGNFSSQCAFSCSEGTNLGTIEFTTGGPFGMWSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCCHPLASFSTGACPFICSEGTLLGKKTTCSSGIMNSPSPIC 120
DB 257 IQCEPLAPDLGIMNCCHPLASFSTGACPFICSEGTLLGKKTTCSSGIMNSPSPIC 316
QY 121 QKLDKSFMIKE 132
DB 317 QKLDKSFMIKE 328

RESULT 2
LEMI_PANTR ID LEMI_PANTR STANDARD; PRT; 372 AA.
AC 095237;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE molecule-1) (LAW-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (IECAM1) (CD62L).
GN SELL.
OS Pan Troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsurushita N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBD databases.
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENUES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U73728; AAB18248.1; -
 DR HSSP: P14151; IKB.1.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_g_1.
 DR Pfam: PF00084; sushi_2.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP_2.
 DR SMART: SM00034; CLECT_1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS00615; C-type-LECTIN_1; 1.
 DR PROSITE: PS00615; C-type-LECTIN_2; 1.
 DR PROSITE: PS00615; C-type-LECTIN_2; 1.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; signal; Sushi; Repeat.
 FT SIGNAL 1 28
 FT PROPEP 29 38
 FT CHAIN 39 372
 FT DOMAIN 39 372
 FT TRANSMEM 333 355
 FT DOMAIN 356 372
 FT DOMAIN 55 155
 FT DOMAIN 156 192
 FT DOMAIN 196 253
 FT DOMAIN 258 317
 FT DISULFID 57 155
 FT DISULFID 128 147
 FT DISULFID 160 171
 FT DISULFID 165 180
 FT DISULFID 182 191
 FT DISULFID 197 241
 FT DISULFID 227 254
 FT DISULFID 259 303
 FT DISULFID 289 316
 FT CARBOHYD 60 60
 FT CARBOHYD 104 104
 FT CARBOHYD 177 177
 FT CARBOHYD 216 216
 FT CARBOHYD 232 232
 FT CARBOHYD 246 246
 FT CARBOHYD 271 271
 SQ SEQUENCE 372 AA; 42188 MW; 6EA991802A2D3643 CRC64;

Query Match 99.2%; Score 727; DB 1; Length 372;
 Best Local Similarity 99.2%; Pred. No. 7.5e-64;
 Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CEPLAPELGTMDCTHHPGNGFSFSCAFSCSEGTNLGIEETTCGPFNGMSPEPTCOV 60
 DB 197 CEPLAPELGTMDCTHHPGNGFSFSCAFSCSEGTNLGIEETTCGPFNGMSPEPTCOV 256
 OY 61 IOCEPLASAPDLGIMNCSPHPLASFSFTSACTFICSEGTTELIGKKKTKICSSGIWNSPSPIC 120
 DB 257 IOCEPLASAPDLGIMNCSPHPLASFSFTSACTFICSEGTTELIGKKKTKICSSGIWNSPSPIC 316
 OY 121 OKLDSFSMIKE 132
 DB 317 OKLDSFSMIKE 328

RESULT 3
 LEM1_PONPY

ID LEM1_PONPY STANDARD; PRT; 372 AA.
 AC 095235;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion
 DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
 DE (LECAM1) (CD62L).
 GN SEL.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
 RA Tsurushita N.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
 CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
 CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
 CC VENULES IN PERIPHERAL LYMPH NODES.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U73729; AAB18247.1; -
 DR HSSP: P14151; IKB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_g_1.
 DR Pfam: PF00084; sushi_2.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP_2.
 DR SMART: SM00034; CLECT_1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS00615; C-type-LECTIN_1; 1.
 DR PROSITE: PS00615; C-type-LECTIN_2; 1.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; signal; Sushi; Repeat.
 FT SIGNAL 1 28
 FT PROPEP 29 38
 FT CHAIN 39 372
 FT DOMAIN 39 372
 FT TRANSMEM 333 355
 FT DOMAIN 356 372
 FT DOMAIN 55 155
 FT DOMAIN 156 192
 FT DOMAIN 196 253
 FT DOMAIN 258 317
 FT DISULFID 57 155
 FT DISULFID 128 147
 FT DISULFID 160 171
 FT DISULFID 165 180
 FT DISULFID 182 191
 FT DISULFID 197 241
 FT DISULFID 227 254
 FT DISULFID 259 303
 FT DISULFID 289 316

```

FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 42118 MW; 6517DD22213FE15E CRC64;

Query Match 98.8%; Score 724; DB 1; Length 372;
Best Local Similarity 98.5%; Pred. No. 1.5e-63;
Matches 130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSSQCAFSCSGTNTLGTETTCGPGNWSPEPTCOV 60
DB 197 CEPLAPDLGMDCTHPGNGFSSQCAFSCSGTNTLGTETTCGPGNWSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTICSSGIMSNPSPIC 120
DB 257 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTICSSGIMSNPSPIC 316
QY 121 QKDKSFSMIKE 132
DB 317 QKDKSFSMIKE 328

RESULT 4
LEML_MACMU STANDARD; PRT; 372 AA.
ID LEML_MACMU
AC Q95198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (leukocyte adhesion molecule-1) (LAM-1) (leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
DE GN
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L., Tsunashita N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U73730; AB18246.1; -
DR HSSP: P14151; IAKJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; Egf_1.

```

```

DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sush1; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sush1; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
FT DISULFID 128 147
FT DISULFID 160 171
FT DISULFID 165 180
FT DISULFID 182 191
FT DISULFID 197 241
FT DISULFID 227 254
FT DISULFID 259 303
FT DISULFID 289 316
FT CARBOHYD 60 60
FT CARBOHYD 104 104
FT CARBOHYD 177 177
FT CARBOHYD 226 226
FT CARBOHYD 232 232
FT CARBOHYD 246 246
FT CARBOHYD 271 271
SQ SEQUENCE 372 AA; 42109 MW; 08BD7DD5AC549D6D CRC64;

Query Match 94.1%; Score 690; DB 1; Length 372;
Best Local Similarity 92.4%; Pred. No. 3e-60;
Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSSQCAFSCSGTNTLGTETTCGPGNWSPEPTCOV 60
DB 197 CEPLAPDLGMDCTHPGNGFSSQCAFSCSGTNTLGTETTCGPGNWSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTICSSGIMSNPSPIC 120
DB 257 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTICSSGIMSNPSPIC 316
QY 121 QKDKSFSMIKE 132
DB 317 QKDKSFSMIKE 328

RESULT 5
LEML_PAPHA STANDARD; PRT; 372 AA.
ID LEML_PAPHA
AC Q28768;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (leukocyte adhesion molecule-1) (LAM-1) (leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
DE GN
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Papio.

```

OX NCBI_TaxID=9557;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97128794; PubMed-8973334;
 RA Tsunashita N., Fu H., Berg E.L.;
 RT "PCR cloning of the cDNA encoding baboon L-selectin.";
 RL Gene 181:219-220(1996).
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE
 CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
 CC VENULES IN PERIPHERAL LYMPH NODES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U52074; AAB40903.1; -
 CC HSSP: P14151; 1KJB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sushi_2.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP; 2.
 DR SMART: SM00181; EGF; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C_Type_Lectin_1; 1.
 DR PROSITE: PS50041; C_Type_Lectin_2; 1.
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sushi; Repeat.
 FT SIGNAL 1 28
 FT PROPEP 29 38
 FT CHAIN 39 372
 FT DOMAIN 39 332
 FT TRANSMEM 333 355
 FT DOMAIN 356 372
 FT DOMAIN 55 155
 FT DOMAIN 156 192
 FT DOMAIN 196 255
 FT DOMAIN 258 317
 FT DOMAIN 317 355
 FT DISULFID 57 155
 FT DISULFID 128 147
 FT DISULFID 160 171
 FT DISULFID 165 180
 FT DISULFID 182 191
 FT DISULFID 197 241
 FT DISULFID 227 254
 FT DISULFID 259 303
 FT DISULFID 289 316
 FT CARBOHYD 60 104
 FT CARBOHYD 104 177
 FT CARBOHYD 177 226
 FT CARBOHYD 226 232
 FT CARBOHYD 232 246
 FT CARBOHYD 246 271
 FT CARBOHYD 271 372
 FT SEQUENCE 372 AA; 42091 MW; 64E7BD5AC549D69 CRC64;

Query Match

94.1%; Score 690; DB 1; Length 372;

Best Local Similarity 92.4%; Pred. No. 3e-60;
 Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CEPLAEPDLCTMDCTHPGNGFSSQCAFSCEGTNLGTIEETTCGPGFNGWSSPEPTQCV 60
 Db 197 CEPLAEPDLCTMDCTHPGNGFSSQCAFSCEGTNLGTIEETTCGPGFNGWSSPEPTQCV 256
 QY 61 IOCEPLAEPDLGIMNCNSHPLASFSTFSACTFCSSEGTLELGGKRTCESSGIMSNPSPIC 120
 Db 257 IOCEPLAEPDLGIMNCNSHPLASFSTFSACTFCSSEGTLELGGKRTCESSGIMSNPSPIC 316
 QY 121 QKDRSFSMIKE 132
 Db 317 QKDRSFSMIKE 328
 RESULT 6
 ID LEM1_BOVIN STANDARD; PRT; 370 AA.
 AC P98131;
 DT 01-FEB-1996 (rel. 33, Created)
 DT 01-FEB-1996 (rel. 33, Last sequence update)
 DT 01-FEB-1996 (rel. 33, Last annotation update)
 DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
 DE SELL.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92164727; PubMed-13711468;
 RA Walcheck B., White M., Kurk S., Kishimoto T.K., Jurilla M.A.;
 RT "Characterization of the bovine peripheral lymph node homing
 RT receptor: a lectin cell adhesion molecule (LECAM).";
 RL Eur. J. Immunol. 22:469-476(1992).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94055053; PubMed-7694420;
 RA Bosworth B.T., Dowbenko D., Shuster D.E., Harp J.A.;
 RT "Bovine L-selectin: a peripheral lymphocyte homing receptor.";
 RL Vet. Immunol. Immunopathol. 37:201-215(1993).
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE
 CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
 CC VENULES IN PERIPHERAL LYMPH NODES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X62882; CAA44676.1; -
 CC HSSP: P14151; 1KJB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sushi_2.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP; 2.

DR	SMART; SMO00034; CLECTY_1.
DR	SMART; SMO00181; EGF_1.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; P50116; EGF_2; 1.
DR	PROSITE; P500615; C_Type_Lectin_1; 1.
DR	PROSITE; P550041; C_Type_Lectin_2; 1.
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectina
RK	Selectin; Signal; Sushi; Repeat.
FT	SIGNAL 1 28
FT	PROPEP 29 38
FT	CHAIN 39 370
FT	DOMAIN 39 333
FT	TRANSMEM 334 354
FT	DOMAIN 355 370
FT	DOMAIN 55 155
FT	DOMAIN 156 192
FT	DOMAIN 196 255
FT	DOMAIN 258 317
FT	DISEUFIID 57 155
FT	DISEUFIID 126 147
FT	DISEUFIID 160 171
FT	DISEUFIID 165 180
FT	DISEUFIID 182 191
FT	DISEUFIID 197 241
FT	DISEUFIID 227 254
FT	DISEUFIID 259 303
FT	DISEUFIID 289 316
FT	CARBOHYD 60 60
FT	CARBOHYD 77 77
FT	CARBOHYD 104 104
FT	CARBOHYD 177 177
FT	CARBOHYD 216 216
FT	CARBOHYD 226 226
FT	CARBOHYD 246 246
FT	CARBOHYD 308 308
FT	CARBOHYD 320 320
FO	SEQUENCE 370 AA; 4197 MW; 92168f8b116ae9228 CRC64;

Query Match	73.3%	Score 537	DB 1	Length 370
Best Local Similarity	68.9%	Pred. No. 2	5e-45	
Matches 91	Conservative 19	Mismatches 22	Indels 0	Gaps
QY	1	CEPLAPELGTMDCTHPPGNFSFSSQCAFSCSEGTNLTGIEETTCGPGNMSSPEPTQY	60	
Db	197	CVLEELPGLGTMACTHPLGNFSFMSQCAFNCSTGMDMGVEETTCAPFGNMSSPEPTCR	256	
QY	61	IOCEPLASPDGIGMNCSPPLASFSFTSACTCTGCEGTEGLKKTKTESSGIWNSPSPIC	120	
Db	257	IOCEPLTEPDDGTMDCMNHLVDFGSSSTCTPSCSEALTEGKTKTCGLSGNMSSPSPC	310	
QY	121	OKDKSFSMIKE	132	
Db	317	OKINRTISINEE	328	

RESULT	7
LEML_RAT	
ID	STANDARD
AC	P30836
DT	01-JUL-1993 (Rel. 26, Created)
DT	01-JUL-1993 (Rel. 26, Last sequence update)
DT	01-FEB-1996 (Rel. 33, Last annotation update)
DE	L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (LY-22) (Lymphocyte surface MEL-14 antigen)
DE	(Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN	SEL1 OR LNH1 OR LY-22.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX	NCBI_TaxId=10116;
RN	[1]
RP	SEQUENCE FROM N.A.

RA	MEDLINE=92329548; PubMed=1378303;
RX	Matanabe T., Song Y., Hirayama Y., Tamatani T., Kuida K., Miyasaka M.;
RT	"Sequence and expression of a rat cDNA for LECAM-1.";
RL	Biochim. Biophys. Acta 1131:321-324(1992).
CC	-!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC	-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC	-!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC	-!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; D10831; BAA01613.1; .
DR	PIR; S23936; S23936.
DR	HSSP; P14151; IKTB.
DR	InterPro; IPR000561; EGF-like.
DR	InterPro; IPR002396; Selectin.
DR	InterPro; IPR000436; Sushi_SCR_CCP.
DR	InterPro; IPR001304; lectln_c.
DR	Pfam; PF00008; EGF_1.
DR	Pfam; PF00059; lectln_c; 1.
DR	Pfam; PF00084; sushi_2.
DR	PRINTS; PR00343; SELECTIN.
DR	SMART; SM00032; CCP; 2.
DR	SMART; SM00034; CLECT; 1.
DR	SMART; SM00181; EGF_1.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS00615; C_type_Lectin_1; 1.
DR	PROSITE; PS00041; C_type_Lectin_2; 1.
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW	Selectin; Signal; Sushi; Repeat.
FT	SIGNAL 1 28
FT	PROPEP 39 38
FT	CHAIN 39 372
FT	DOMAIN 39 332
FT	TRANSMEM 333 355
FT	DOMAIN 356 372
FT	DOMAIN 55 155
FT	DOMAIN 156 192
FT	DOMAIN 196 255
FT	DOMAIN 258 317
FT	DISULFID 57 155
FT	DISULFID 128 147
FT	DISULFID 160 171
FT	DISULFID 165 180
FT	DISULFID 182 191
FT	DISULFID 197 241
FT	DISULFID 227 254
FT	DISULFID 259 303
FT	DISULFID 289 316
FT	CARBOHYD 60 60
FT	CARBOHYD 104 104
FT	CARBOHYD 177 177
FT	CARBOHYD 226 226
FT	CARBOHYD 246 246
FT	CARBOHYD 278 278
SEQUENCE	372 AA; 42441 MW; 3B8BAE0F154D191A CRC64;

Query Match	73.0%;	Score 535;	DB 1;	Length 372;
Best Local Similarity	71.2%;	Pred. No. 4e-45;		
Matches	94;	Conservative	15;	Mismatches 23;
			Indels	0;
			Gaps	0;


```
OY 1 CEPLPAPDLGIMNCSDHPLASFSFTSACTFTCSGTELLIGKRTICSSGIGWNSPSPIC 60
DB 197 CEPLPAPDLGIMNCSDHPLASFSFTSACTFTCSGTELLIGKRTICSSGIGWNSPSPIC 256
OY 61 IQCEPLPAPDLGIMNCSDHPLASFSFTSACTFTCSGTELLIGKRTICSSGIGWNSPSPIC 120
DB 257 IQCEPLPAPDLGIMNCSDHPLASFSFTSACTFTCSGTELLIGKRTICSSGIGWNSPSPIC 316
OY 121 OKLDSFSMIKE 132
DB 317 OKTRRSFSKIKE 328

RESULT 8
LEML_MOUSE STANDARD: PRT: 372 AA.
AC P18337;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (LY-22) (lymphocyte surface MEL-14 antigen) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN SELL OR LNHOR OR LY-22.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Siegelman M.H., van de Rijm M., Weissman I.L.;
RX MEDLINE=89162048; PubMed=2646713;
RT "Mouse lymph node homing receptor cDNA clone encodes a glycoprotein revealing tandem interaction domains.";
RL Science 243:1165-1172(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90263086; PubMed=1693096;
RA Siegelman M.H., Cheng I.C., Weissman I.L., Wakeland E.K.;
RT "The mouse lymph node homing receptor is identical with the lymphocyte cell surface marker Ly-22: role of the EGF domain in endothelial binding.";
RL Cell 61:611-622(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89168433; PubMed=2647302;
RA Laaky L.A., Singer M.S., Yednock T.A., Dowbenko D., Fennie C., Rodriguez H., Nguyen T., Stachel S., Rosen S.D.;
RT "Cloning of a lymphocyte homing receptor reveals a lectin domain.";
RL Cell 56:1045-1055(1989).
RN [4]
RP SEQUENCE OF 1-360 FROM N.A.
RX MEDLINE=91169529; PubMed=2004776;
RA Dowbenko D.J., Diep A., Taylor B.A., Lusis A.J., Lasky L.A.;
RT "Characterization of the murine homing receptor gene reveals correspondence between protein domains and coding exons.";
RL Genomics 9:270-277(1991).
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECTINS.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
```

```
CC -----
DR EMBL: X14772; CAA32880.1; -
DR EMBL: M36005; AAA39722.1; -
DR EMBL: M36058; AAA39723.1; -
DR EMBL: M25324; AAA39431.1; -
DR EMBL: M64549; AAA75651.1; -
DR EMBL: M64440; AAA75651.1; JOINED.
DR EMBL: M64545; AAA75651.1; JOINED.
DR EMBL: M64548; AAA75651.1; JOINED.
DR PIR: A32375; A32375.
DR HSSP: P14151; 1KCB.
DR MGD: MGI:98279; Sell.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR001212; Somatomedin_B.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c.
DR Pfam: PF00084; sushi_2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00201; SO; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-type-LECTIN_1; 1.
DR PROSITE: PS50041; C-type-LECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFD 57 155
FT DISULFD 128 147
FT DISULFD 160 171
FT DISULFD 165 180
FT DISULFD 182 191
FT DISULFD 197 241
FT DISULFD 227 254
FT DISULFD 259 303
FT DISULFD 289 316
FT CARBOHYD 60 60
FT CARBOHYD 104 104
FT CARBOHYD 176 177
FT CARBOHYD 216 216
FT CARBOHYD 226 226
FT CARBOHYD 246 246
FT CARBOHYD 278 278
FT CARBOHYD 308 308
FT CARBOHYD 320 320
FT CARBOHYD 320 320
FT CONFLICT 32 32
FT SEQUENCE 372 AA; 42288 MW; 4433EDF6ACB2B78 CMC64;
Query Match 70.7%; Score 518; DB 1; Length 372;
Best Local Similarity 68.2%; Pred. No. 1.8e-43;
Matches 90; Conservative 13; Mismatches 29; Indels 0; Gaps 0;
OY 1 CEPLPAPDLGIMNCSDHPLASFSFTSACTFTCSGTELLIGKRTICSSGIGWNSPSPIC 60
DB 197 CEPLPAPDLGIMNCSDHPLASFSFTSACTFTCSGTELLIGKRTICSSGIGWNSPSPIC 256
OY 61 IQCEPLPAPDLGIMNCSDHPLASFSFTSACTFTCSGTELLIGKRTICSSGIGWNSPSPIC 120
```

```

Db      257 VQCEPLEAPFELGTMDCHPLGNFSFQSCAPNCSEGBELLGTARTQCGASGNWSPSPIC 316
Oy      121 QKDKSFSMAKE 132
       1 : : : : :
Db      317 QETNRSPFSKIKE 328

RESULT 9
LEK3_HUMAN
ID      LEM3_HUMAN          STANDARD;          PRT;          830 AA.
AC      P16109;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
DE      (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN      SLIP OR GMRP.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89168432; Pubmed=2466574;
RA      Johnston G.I., Cook R.G., McEver R.P.;
RT      Cloning of GMP-140, a granule membrane protein of platelets and
RT      endothelium: sequence similarity to proteins involved in cell
RT      adhesion and inflammation."
RL      Cell 56:1033-1044(1989).
RN      [2]
RP      PALMITOYLATION.
RX      MEDLINE=93266599; Pubmed=7684381;
RA      Fujimoto T., Stroud E., Whately R.E., Prescott S.M., Muszbek L.,
RA      Laposta M., McEver R.P.;
RT      P-selectin is acylated with palmitic acid and stearic acid at
RT      cysteine 766 through a thioester linkage."
RL      J. Biol. Chem. 268:11394-11400(1993).
RN      [3]
RP      STRUCTURE BY NMR OF 160-199.
RX      MEDLINE=97057176; Pubmed=8901515;
RA      Freedman S.J., Sanford D.G., Bachovchin W.W., Furie B.C., Baleja J.D.,
RA      Furie B.;
RT      Structure and function of the epidermal growth factor domain of P-
RT      selectin."
RL      Biochemistry 35:13733-13744(1996).
RN      [4]
RP      3D-STRUCTURE MODELING OF 42-161.
RX      MEDLINE=94093388; Pubmed=7505680;
RA      Bajorek J., Stenkamp R., Arnifio A.;
RT      "Knowledge-based model building of proteins: concepts and examples.";
RT      Protein Sci. 2:1798-1810(1993).
RN      [5]
RP      VARIANTS ASN-331; ASP-603; VAL-640 AND PRO-756.
RX      MEDLINE=98334547; Pubmed=9668170;
RA      Hermann S.M., Ricard S., Nicaud V., Mallet C., Evans A.,
RA      Rudevets J.B., Arweiler D., Luc G., Cambien F.;
RT      "The P-selectin gene is highly polymorphic: reduced frequency of the
RT      pro715 allele carriers in patients with myocardial infarction.";
RL      Hum. Mol. Genet. 7:1277-1284(1998).
CC      -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS
CC      TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC      INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC      LEUKOCYTES. THE LIGAND RECOGNIZED IS STAYL-Lewis X.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
CC      AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
CC      ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
CC      THE CELL SURFACE.
CC      -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC      -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      -1- SIMILARITY: CONTAINS 9 SUSHI (SCR) DOMAINS.

```

```

CC      -1- DATABASE: NAME=PROW; NOTE=CD guide CD62P entry;
CC      WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62p.htm".
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      of the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@isb-slb.ch).
CC      -----
DR      EMBL; M60234; AAA35910.1; -
DR      EMBL; M60217; AAA35910.1; JOINED.
DR      EMBL; M60218; AAA35910.1; JOINED.
DR      EMBL; M60219; AAA35910.1; JOINED.
DR      EMBL; M60222; AAA35910.1; JOINED.
DR      EMBL; M60223; AAA35910.1; JOINED.
DR      EMBL; M60224; AAA35910.1; JOINED.
DR      EMBL; M60225; AAA35910.1; JOINED.
DR      EMBL; M60226; AAA35910.1; JOINED.
DR      EMBL; M60227; AAA35910.1; JOINED.
DR      EMBL; M60228; AAA35910.1; JOINED.
DR      EMBL; M60229; AAA35910.1; JOINED.
DR      EMBL; M60231; AAA35910.1; JOINED.
DR      EMBL; M60232; AAA35910.1; JOINED.
DR      EMBL; M60233; AAA35910.1; JOINED.
DR      EMBL; M25322; AAA35911.1; -
DR      PIR; A30359; A30359.
DR      PDB; 1FSB; 01-APR-97.
DR      PDB; 1KDD; 03-APR-96.
DR      MIM; 173610; -
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR002396; Selectin.
DR      InterPro; IPR000436; Sushi_SCR_CCP.
DR      InterPro; IPR001304; Lectin_C.
DR      Pfam; PF00008; EGF_1.
DR      Pfam; PF00059; Lectin_C; 1.
DR      Pfam; PF00084; sushi; 9.
DR      PRINTS; PR00343; SELECTIN.
DR      SMART; SM00032; CCP; 9.
DR      SMART; SM00034; CLECT; 1.
DR      SMART; SM00181; EGF_1.
DR      PROSITE; PS00022; EGF_1; 1.
DR      PROSITE; PS00186; EGF_2; 1.
DR      PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
DR      PROSITE; PS50041; C-TYPE_LLECTIN_2; 1.
KW      Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW      Selectin; Signal; Sushi; Repeat; Lipoprotein; Palmitate; Polymorphism;
KW      3D-structure.
FT      SIGNAL 1 41
FT      CHAIN 42 830
FT      DOMAIN 42 771
FT      TRANSMEM 772 795
FT      DOMAIN 796 830
FT      DOMAIN 158 158
FT      DOMAIN 159 195
FT      DOMAIN 199 258
FT      DOMAIN 261 320
FT      DOMAIN 323 382
FT      DOMAIN 385 444
FT      DOMAIN 447 506
FT      DOMAIN 509 568
FT      DOMAIN 571 630
FT      DOMAIN 641 702
FT      DOMAIN 703 762
FT      DISULFID 60 158
FT      DISULFID 131 150
FT      DISULFID 163 174
FT      DISULFID 168 183
FT      DISULFID 185 194
FT      DISULFID 200 244
FT      DISULFID 230 257
FT      DISULFID 262 306

```

	Query Match Similarity	43.0%	Score 315;	DB 1;	Length 830;
	Best Local Similarity	43.9%	Pred. No. 2,6e-23;		
Matches	54;	Conservative	21;	Mismatches	46;
				Indels	0;
				Gaps	0;
OY	1	CEPLAEPGLIGMCDTHPEGNFSSOCAPSCSEBGTNLGIEETTGGPFGNWSSPETTCOV	60		
		: : : : : :	: : :	:	:
Db	324	CCHLEPAPEGSTGMDCVHPILFAFYAGSGCKECCPGYGVNRDIDMIRCIDSHWASAPLTCEA	383		
OY	61	IOCEPLSADLIGMNCSHPLASFPSFTACTFCSECTELTGKKTKTCESSGIAMSNSPIC	120		
		: : : : : : : :	: : : : : :	:	: : :
Db	364	TICEPLELESYHWSMDCSPSLRAFYDTNCSFRCABEFMLRGADIYCNDLGQMTAFAPVC	443		
OY	121	OKL 123			
Db	444	QAL 446			
 RESULT 10 LEM3_RAT ID_LBM1_RAT STANDARD: PRT: 768 AA. AC P98106; DT 01-FEB-1996 (Rel. 33, Created) DT 01-FEB-1996 (Rel. 33, Last sequence update) 01-MAR-2002 (Rel. 41, Last annotation update) DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM DE (CMB2P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3). GN SEIP. OS Rattus norvegicus (Rat). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. OX NCBI_TaxId=10116; RN [1]					

RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung;
RX	MEDLINE=94333817; PubMed=7520013;
RA	Auchampach J.A., Oliver M.G., Anderson D.C., Manning A.M.;
RT	"Cloning, sequence comparison and in vivo expression of the gene
RL	encoding rat p-selectin..";
CC	Gene 145.251-255(1994)
CC	-1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC	TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC	INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC	LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED: SPLEEN,
CC	LUNG, BRAIN, LIVER, HEART, KIDNEY, THYMUS, SMALL INTESTINE.
CC	-1- INDUCTION: BY ACUTE INFLAMMATION (PROBABLE).
CC	-1- SIMILARITY: TO OTHER SELECTINS/LECTINS.
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC	-1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS; RAT P-LECTIN LACKS THE
CC	HUMAN SUSHI-2 EQUIVALENT.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: L23088; AAB60325.1; -
DR	HSPB: P16109; IERSB.
DR	InterPro: IPR000561; EGF-like.
DR	InterPro: IPR002396; Selectin.
DR	InterPro: IPR00436; Sush1_SCR_CCP.
DR	InterPro: IPR001304; lectin_c.
DR	Pfam: PF00059; lectin_c.1.
DR	Pfam: PF00084; sush1; 8.
DR	PRINTS: PR00343; SELECTIN.
DR	SMART: SM00032; CCP; 8.
DR	SMART: SM00034; CLECT; 1.
DR	SMART: SM00181; EGF; 1.
DR	PROSITE: PS00022; EGF_1; 1.
DR	PROSITE: PS01186; EGF_2; 1.
DR	PROSITE: PS00615; C-type-LECTIN_1; 1.
DR	PROSITE: PS00411; C-type-LECTIN_2; 1.
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW	Selectin; Signal; Sush1; Repeat; Lipoprotein; Palmitate.
KW	SIGNAL
FT	CHAIN 1 41
FT	FT CHAIN 42 768
FT	FT DOMAIN 42 709
FT	FT TRANSMEM 710 733
FT	FT DOMAIN 734 768
FT	FT DOMAIN 58 158
FT	FT DOMAIN 159 195
FT	FT DOMAIN 199 258
FT	FT DOMAIN 261 320
FT	FT DOMAIN 323 382
FT	FT DOMAIN 385 444
FT	FT DOMAIN 447 506
FT	FT DOMAIN 509 568
FT	FT DOMAIN 579 638
FT	FT DOMAIN 641 700
FT	FT DISULFID 60 158
FT	FT DISULFID 131 150
FT	FT DISULFID 168 183
FT	FT DISULFID 185 194
FT	FT DISULFID 200 244
FT	FT DISULFID 230 257
FT	FT DISULFID 262 306
FT	FT DISULFID 292 319
FT	FT DISULFID 324 368
FT	FT DISULFID 354 381
FT	FT DISULFID 386 430
FT	FT DISULFID 386 430
FT	FT SIGNAL 1 41
FT	FT CHAIN 42 768
FT	FT DOMAIN 42 709
FT	FT TRANSMEM 710 733
FT	FT DOMAIN 734 768
FT	FT DOMAIN 58 158
FT	FT DOMAIN 159 195
FT	FT DOMAIN 199 258
FT	FT DOMAIN 261 320
FT	FT DOMAIN 323 382
FT	FT DOMAIN 385 444
FT	FT DOMAIN 447 506
FT	FT DOMAIN 509 568
FT	FT DOMAIN 579 638
FT	FT DOMAIN 641 700
FT	FT DISULFID 60 158
FT	FT DISULFID 131 150
FT	FT DISULFID 168 183
FT	FT DISULFID 185 194
FT	FT DISULFID 200 244
FT	FT DISULFID 230 257
FT	FT DISULFID 262 306
FT	FT DISULFID 292 319
FT	FT DISULFID 324 368
FT	FT DISULFID 354 381
FT	FT DISULFID 386 430
FT	FT DISULFID 386 430
FT	FT SIGNAL 1 41
FT	FT CHAIN 42 768
FT	FT DOMAIN 42 709
FT	FT TRANSMEM 710 733
FT	FT DOMAIN 734 768
FT	FT DOMAIN 58 158
FT	FT DOMAIN 159 195
FT	FT DOMAIN 199 258
FT	FT DOMAIN 261 320
FT	FT DOMAIN 323 382
FT	FT DOMAIN 385 444
FT	FT DOMAIN 447 506
FT	FT DOMAIN 509 568
FT	FT DOMAIN 579 638
FT	FT DOMAIN 641 700
FT	FT DISULFID 60 158
FT	FT DISULFID 131 150
FT	FT DISULFID 168 183
FT	FT DISULFID 185 194
FT	FT DISULFID 200 244
FT	FT DISULFID 230 257
FT	FT DISULFID 262 306
FT	FT DISULFID 292 319
FT	FT DISULFID 324 368
FT	FT DISULFID 354 381
FT	FT DISULFID 386 430
FT	FT DISULFID 386 430
FT	FT SIGNAL 1 41
FT	FT CHAIN 42 768
FT	FT DOMAIN 42 709
FT	FT TRANSMEM 710 733
FT	FT DOMAIN 734 768
FT	FT DOMAIN 58 158
FT	FT DOMAIN 159 195
FT	FT DOMAIN 199 258
FT	FT DOMAIN 261 320
FT	FT DOMAIN 323 382
FT	FT DOMAIN 385 444
FT	FT DOMAIN 447 506
FT	FT DOMAIN 509 568
FT	FT DOMAIN 579

RESULT 13		
ID	LEM3_MOUSE	LEM3_MOUSE
AC	001102:	001102:
DT	01-APR-1993	(rel. 25, Created)
DT	01-APR-1993	(rel. 25, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
 DE (CD62P) (leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
 GN SEMP OR GRMP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92340571; PubMed=1378846;
 RA Weiler A., Iseemann S., Vestweber D.;
 RT Cloning of the mouse endothelial selectins. Expression of both E-
 RT and P-selectin is inducible by tumor necrosis factor alpha.";
 RL J. Biol. Chem. 267:15176-15183(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92345617; PubMed=1379089;
 RA Sanders W.E. Jr., Wilson R.W., Ballantyne C.M., Beaudet A.L.;
 RT Molecular cloning and analysis of in vivo expression of murine P-
 RT selectin.";
 RL Blood 80:795-800(1992).
 CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
 CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
 CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
 CC THE CELL SURFACE.
 CC -1- INDUCTION: BY TNF-ALPHA.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS; MOUSE P-LECTIN LACKS
 CC THE HUMAN SUSHI-2 EQUIVALENT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M87861; AAA40008.1; -;
 DR EMBL: M72332; AAA3712.1; -;
 DR PIR: A42755; A42755.
 DR HSSP: P16109; IFSB.
 DR MGD: MGI:98280; Selp.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF 2.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sush1-SCR_CCP.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR Pfam: PF00084; sush1; 8.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP; 8.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sush1; Repeat; Lipoprotein; Palmitate.
 FT SIGNAL 1 41
 FT CHAIN 42 768 P-SELECTIN.
 FT DOMAIN 42 709 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 710 733 POTENTIAL.

FT DOMAIN 734 768 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 58 158 C-TYPE LECTIN (SHORT FORM).
 FT DOMAIN 159 195 EGF-LIKE.
 FT DOMAIN 199 258 SUSHI 1.
 FT DOMAIN 261 320 SUSHI 2.
 FT DOMAIN 323 382 SUSHI 3.
 FT DOMAIN 385 444 SUSHI 4.
 FT DOMAIN 447 506 SUSHI 5.
 FT DOMAIN 509 568 SUSHI 6.
 FT DOMAIN 579 638 SUSHI 7.
 FT DOMAIN 641 700 SUSHI 8.
 FT DISULFID 60 158 BY SIMILARITY.
 FT DISULFID 131 150 BY SIMILARITY.
 FT DISULFID 163 174 BY SIMILARITY.
 FT DISULFID 168 183 BY SIMILARITY.
 FT DISULFID 185 194 BY SIMILARITY.
 FT DISULFID 200 244 BY SIMILARITY.
 FT DISULFID 230 257 BY SIMILARITY.
 FT DISULFID 262 306 BY SIMILARITY.
 FT DISULFID 292 319 BY SIMILARITY.
 FT DISULFID 324 368 BY SIMILARITY.
 FT DISULFID 354 381 BY SIMILARITY.
 FT DISULFID 386 430 BY SIMILARITY.
 FT DISULFID 416 443 BY SIMILARITY.
 FT DISULFID 448 492 BY SIMILARITY.
 FT DISULFID 478 505 BY SIMILARITY.
 FT DISULFID 510 554 BY SIMILARITY.
 FT DISULFID 540 567 BY SIMILARITY.
 FT DISULFID 580 624 BY SIMILARITY.
 FT DISULFID 610 637 BY SIMILARITY.
 FT DISULFID 642 686 BY SIMILARITY.
 FT DISULFID 672 699 BY SIMILARITY.
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 654 654 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 661 661 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT LIPID 745 745 PALMITATE (BY SIMILARITY).
 FT SITE 756 759 ENDOCYTOSIS SIGNAL (PROBABLE).
 FT CONFLICT 724 724 A -> E (IN REF. 2).
 SQ SEQUENCE 768 AA; 83098 MW; E5173074D2P6568 CRC64;

 Query Match 40.9%; Score 300; DB 1; Length 768;
 Best local Similarity 42.3%; Pred. No. 6; 9e-22;
 Matches 52; Conservative 17; Mismatches 54; Indels 0; Gaps 0;

 QY 1 CEPLAPDLGTMCTHFGNPSRSSQCAFSCSESTNTGTIEFTTCGPFQWSSPEPTCOV 60
 DB 200 CGKVNIPQHVLMNCSHPLGESENSQCTFSCABEYELDGELOCLASGIWNNPPKCD 259
 QY 61 IOCEPLSAPDLGIMNCNSHPLASFSTACTFICSEGTETIGKKKTCSSGIMNSPDC 120
 DB 260 VQCCSLAPHPHGMACHMPLAFAVYDSSCKFEQCPGRANGSNMTHOTSGGQNSPLPTC 319
 QY 121 QKL 123
 DB 320 EAI 322

 RESULT 14
 LEM3_SHEEP STANDARD; PRT; 769 AA.
 ID LEM3_SHEEP
 AC P98109;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
 DE (CD62P) (leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
 GN SEMP.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

CC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=99940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RA Burns S.A., Neufeld E.J., Donady J.J.;
 RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
 CC -1 FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
 TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
 INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
 LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1 SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L34270; AAB59261.1; -.
 DR HSSB; P16109; IESB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR InterPro: IPR001304; Lectin_c.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00059; Lectin_c; 1.
 DR Pfam; PF00084; Sush1_8.
 DR PRINTS; PR00343; SELECTIN.
 DR SMART; SM00032; CCP; 8.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF_1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
 DR PROSITE; PS00615; C-TYPE_LECTIN_2; 1.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sush1; Repeat
 FT SIGNAL 1 32 POTENTIAL.
 FT CHAIN 33 769 P-SELECTIN.
 FT DOMAIN 33 717 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 718 734 POTENTIAL.
 FT DOMAIN 735 769 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 58 158 C-TYPE LECTIN.
 FT DOMAIN 159 195 EGF-LIKE.
 FT DOMAIN 199 258 SUSHI 1.
 FT DOMAIN 261 320 SUSHI 2.
 FT DOMAIN 323 382 SUSHI 3.
 FT DOMAIN 385 444 SUSHI 4.
 FT DOMAIN 447 506 SUSHI 5.
 FT DOMAIN 509 568 SUSHI 6.
 FT DOMAIN 580 639 SUSHI 7.
 FT DOMAIN 642 701 SUSHI 8.
 FT DISULFID 60 158 BY SIMILARITY.
 FT DISULFID 131 150 BY SIMILARITY.
 FT DISULFID 163 174 BY SIMILARITY.
 FT DISULFID 168 183 BY SIMILARITY.
 FT DISULFID 185 194 BY SIMILARITY.
 FT DISULFID 200 244 BY SIMILARITY.
 FT DISULFID 230 257 BY SIMILARITY.
 FT DISULFID 262 306 BY SIMILARITY.
 FT DISULFID 292 319 BY SIMILARITY.
 FT DISULFID 324 368 BY SIMILARITY.
 FT DISULFID 354 381 BY SIMILARITY.
 FT DISULFID 386 430 BY SIMILARITY.
 FT DISULFID 416 443 BY SIMILARITY.
 FT DISULFID 448 492 BY SIMILARITY.

FT DISULFID 478 505 BY SIMILARITY.
 FT DISULFID 510 554 BY SIMILARITY.
 FT DISULFID 540 567 BY SIMILARITY.
 FT DISULFID 581 625 BY SIMILARITY.
 FT DISULFID 611 638 BY SIMILARITY.
 FT DISULFID 643 687 BY SIMILARITY.
 FT DISULFID 673 700 BY SIMILARITY.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 680 680 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 757 760 ENDOCYTOSIS SIGNAL (PROBABLE).
 FT VARIANT 566 566 S -> T.
 FT VARIANT 579 579 L -> V.
 SQ SEQUENCE 769 AA; 84317 MW; 23E42575060FAB15 CRC64;
 Query Match 40.9%; Score 300; DB 1; Length 769;
 Best Local Similarity 39.8%; Pred. No. 6,9e-22;
 Matches 49; Conservative 24; Mismatches 30; Indels 0; Gaps 0;
 QY 1 CEPLAPELLGTMDCTHPGPFSSQCAFSCEGTNLNGIEFTGCPFGNWSPEPTCOV 60
 DB 200 CSEFDPQVNVHNCSPRLNFGFKSCSFHCAAGVALNPRELECLASGIWNSPQCV 259
 QY 61 IOCEPLSPADLIGMNCSPHLSAFSTACTFICSBCTELIGKKTKICSSGIMNSPIC 120
 DB 260 VOCPALKSPQCSMSCFHSKAFQHOSSCFCEGFTLVGEVVTALGWTAPTPVC 319
 QY 121 OKL 123
 DB 320 KAI 322
 RESULT 15
 LEM2_RAT
 ID LEM2_RAT STANDARD; PRT; 549 AA.
 AC P98105.
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
 DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
 DE (CD62E).
 GN SELE OR ELAM-1.
 GN Rattus norvegicus (Rat).
 OS Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
 RL Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.
 CC -1 FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
 POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 GLYCOPOLIDS).
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1 SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL: L25527; AAA41113.1; -.
DR HSSP: P15681; 1KJA.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushi; 5.
DR SMART: SM00032; CCP; 5.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; lectin;
KW Selection; Signal; Sushi; Repeat
FT SIGNAL 1 21
FT CHAIN 22 549
FT DOMAIN 22 549
FT TRANSMEM 495 516
FT DOMAIN 517 549
FT DOMAIN 38 138
FT DOMAIN 139 175
FT DOMAIN 179 239
FT DOMAIN 242 301
FT DOMAIN 304 364
FT DOMAIN 367 427
FT DOMAIN 430 486
FT DISULFID 40 138
FT DISULFID 111 130
FT DISULFID 143 154
FT DISULFID 148 163
FT DISULFID 165 174
FT DISULFID 180 225
FT DISULFID 210 238
FT DISULFID 243 287
FT DISULFID 273 300
FT DISULFID 305 350
FT DISULFID 336 363
FT DISULFID 368 413
FT DISULFID 399 426
FT DISULFID 431 472
FT DISULFID 458 485
FT CARBOHYD 25 25
FT CARBOHYD 60 60
FT CARBOHYD 145 145
FT CARBOHYD 192 192
FT CARBOHYD 203 203
FT CARBOHYD 266 266
FT CARBOHYD 313 313
FT CARBOHYD 320 320
FT CARBOHYD 330 333
FT CARBOHYD 441 441
FT CARBOHYD 465 465
SQ SEQUENCE 549 AA: 60079 MW: 85CECHDB7B0144C8 CRC64;

Query Match 40.98; Score 299.5; DB 1; Length 549;
Best Local Similarity 41.08; Pred. No. 5,5e-22;
Matches 50; Conservative 23; Mismatch 48; Indels 1; Gaps 1;

OY 1 CEPLPELGLMCTDCTHPGNGFSSOCAGFCSSGTLNLTGIEETT-CGPFQNWSSPEPTCQ 59
DB 180 COBOETPDHGLMCTDCTHPGNGFSSOCAGFCSSGTLNLTGIEETT-CGPFQNWSSPEPTCQ 59

OY 60 VIOCEPLSAPDLGIMNCSHPLASFSTISACTFISGTELLIGKKRTICSSGIWNSPST 119
DB 240 VVECKALTPAHGVRKSSNNGSYPMWNTCTFDCDEGVRVGAQNLQCTSSGWDNEKPS 259
OY 120 CK 121
DB 300 CK 301

RESULT 16
LENN_HUMAN
ID LEEN2_HUMAN STANDARD; PRT; 610 AA.
AC P16581; P16111;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE OR ELAM1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175359; PubMed=1689848;
RA Hession C., Osborn L., Goff D., Chl-Rosso G., Vassallo C.,
RA Pasak M., Pittack C., Tizard R., Goelz S., McCarthy K., Hopple S.,
RA Lobb R.,
RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning
RT and functional interactions."
RL Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89162047; PubMed=2466335;
RA Bevilacqua M.P., Stengelin S., Gimbrone M.A. Jr., Seed B.,
RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for
RT neutrophils related to complement regulatory proteins and lectins."
RL Science 243:1160-1165(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91115870; PubMed=1703529;
RA Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T.,
RA Gimbrone M.A. Jr., Bevilacqua M.P.,
RT "Structure and chromosomal location of the gene for endothelial-
RT leukocyte adhesion molecule 1."
RL J. Biol. Chem. 266:2466-2473(1991).
RN [4]
RP LIGAND.
RX MEDLINE=91068005; PubMed=1701274;
RA Phillips M.L., Nudelma E., Gaeta F.C., Perez M., Singhal A.K.,
RA Hakomori S., Paulson J.C.,
RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate
RT ligand, sialyl-Lex."
RL Science 250:1130-1132(1990).
RN [5]
RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.
RX MEDLINE=93202275; PubMed=7681016;
RA Mills A.,
RT "Modeling the carbohydrate recognition domain of human E-selectin."
RL FEBS Lett. 319:5-11(1993).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.
RX MEDLINE=94150646; PubMed=7509040;
RA Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S.,
RA Huang K.-S., Presky D.H., Familletti P.C., Wollitzky B.A., Burns D.R.,
RT "Insight into E-selectin/ligand interaction from the crystal
RT structure and mutagenesis of the Lec/EGF domains."
RL Nature 367:532-538(1994).
RN [7]
RP VARIANT ARG-149.
RX MEDLINE=95179107; PubMed=7533025;

RA Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S.,
RA Schulte K.L., Glaser C., Rohde K., Baumann G., Speer A.;
RT "E-selectin polymorphism and atherosclerosis: an association study.";
RL Hum. Mol. Genet. 3:1935-1937(1994).
RN [8]
RP VARIANT ARG-149.
RX MEDLINE-99134508; PubMed-9933738;
RA Ye S.-Q., Usher D., Virgall D., Zhang L.-Q., Yochim S.E., Gupta R.;
RT "A p51 polymorphism detects the mutation of serine-128 to arginine in
RT CD 62E gene - a risk factor for coronary artery disease.";
RL J. Biomed. Sci. 6:18-21(1999).
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOPOLYIDS).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A
CC HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY
CC HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH
CC ANGIOGRAPHICALLY PROVEN SEVERE ATHEROSCLEROSIS COMPARED WITH AN
CC UNSELECTED POPULATION (SER-149).
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
CC -1- DATABASE: NAME-PROW; NOTE-CD guide CD62E entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/CD/cd62e.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M30640; AAA52377.1; -;
DR EMBL: M61893; AAA52375.1; -;
DR EMBL: M61885; AAA52375.1; JOINED.
DR EMBL: M61887; AAA52375.1; JOINED.
DR EMBL: M61888; AAA52375.1; JOINED.
DR EMBL: M61890; AAA52375.1; JOINED.
DR EMBL: M61891; AAA52375.1; JOINED.
DR EMBL: M61892; AAA52375.1; JOINED.
DR EMBL: M24736; AAA52376.1; -;
DR PIR: A32606; A32606.
DR PIR: A35046; A35046.
DR PIR: A38615; A38615.
DR PDB: 1ESL; 3I-AUG-94.
DR PDB: 1KTA; 03-APR-96.
DR MIM: 131210; -;
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; Lectin_c; 1.
DR Pfam: PF00084; sushi; 6.
DR PRINTS: PRO0343; SELECTIN.
DR SMART: SM00032; CCP; 6.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00186; EGF_2; 1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS00615; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS00615; C_TYPE_LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat; Polymorphism; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 1 610
FT DOMAIN 22 556
FT TRANSMEM 557 578
FT E-SELECTIN.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.

FT	DOMAIN	579	610	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	38	138	C-TYPE LECTIN (SHORT FORM).
FT	DOMAIN	139	175	EGF-LIKE.
FT	DOMAIN	179	238	SUSHI 1.
FT	DOMAIN	241	300	SUSHI 2.
FT	DOMAIN	303	363	SUSHI 3.
FT	DOMAIN	366	426	SUSHI 4.
FT	DOMAIN	429	489	SUSHI 5.
FT	DOMAIN	492	548	SUSHI 6.
FT	DISULFID	40	138	
FT	DISULFID	111	130	
FT	DISULFID	143	154	
FT	DISULFID	148	163	
FT	DISULFID	165	174	
FT	DISULFID	180	224	BY SIMILARITY.
FT	DISULFID	210	237	BY SIMILARITY.
FT	DISULFID	242	286	BY SIMILARITY.
FT	DISULFID	272	299	BY SIMILARITY.
FT	DISULFID	304	349	BY SIMILARITY.
FT	DISULFID	335	362	BY SIMILARITY.
FT	DISULFID	367	412	BY SIMILARITY.
FT	DISULFID	398	425	BY SIMILARITY.
FT	DISULFID	430	475	BY SIMILARITY.
FT	DISULFID	461	488	BY SIMILARITY.
FT	DISULFID	493	534	BY SIMILARITY.
FT	DISULFID	520	547	BY SIMILARITY.
FT	CARBOHYD	25	25	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	145	145	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	179	179	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	203	203	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	265	265	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	312	312	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	503	503	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	527	527	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	130	130	C -> W (IN DBSNP:5360).
FT	VARIANT	149	149	/FTID=VAR_011790.
FT	VARIANT	149	149	S -> R (ASSOCIATED WITH A RISK FACTOR FOR CAD).
FT	VARIANT	295	295	/FTID=VAR_004191.
FT	VARIANT	295	295	E -> K (IN DBSNP:5364).
FT	VARIANT	421	421	/FTID=VAR_011791.
FT	VARIANT	421	421	E -> Q (IN DBSNP:5366).
FT	VARIANT	468	468	H -> Y (IN DBSNP:5368).
FT	VARIANT	575	575	/FTID=VAR_011793.
FT	VARIANT	575	575	L -> F (IN DBSNP:5355).
FT	VARIANT	575	575	/FTID=VAR_011794.
FT	SEQUENCE	610 AA; 66655 KM; 7D43E3C0D1229229 CRC64;		

Query Match 40.4%; Score 296; DB 1; Length 610;
Best local Similarity 41.3%; Pred. No. 1,3e-21;
Matches 50; Conservative 21; Mismatch 50; Indels 0; Caps 0;

OY	1	CEPLEAPELGMDCTHPFGNFSFSSQCAFSCEGTNLTGIEETTCGPFGMSSPEPTCOV	60
DB	180	CTALESPEHGSILVCSHPLGNFSYSSCSISCDRGVLPSSMETMQCMSSGEMSAPIPCANV	239
OY	61	IOCEPLSAPDLGIMNCNSHPLASFSFTSACFTISEGELGKKTKTCESGTSINSPJIC	120
DB	240	VECDAVTNPANGVEVECPQNGSFPWNTCTFDEEGELMGASLOCTSSGNDNEKPTC	299
OY	121	Q 121	
DB	300	K 300	

RESULT 17
LEM3_BOVIN
ID LEM3_BOVIN STANDARD; PRT; 646 AA.

AC P42201;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE P-selectin precursor (granule membrane protein 140) (GMP-140) (PADGEM)
 DE (CD62P) (leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
 GN SELEP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Capillary endothelium;
 RA MEDLINE=93249394; PubMed=7683458;
 RT Strubel N.A., Nguyen M., Kansas G.S., Tedder T.F., Bischoff J.;
 RT "Isolation and characterization of a bovine cDNA encoding a
 RT functional homolog of human P-selectin.";
 RL Biochem. Biophys. Res. Commun. 192:338-344(1993).
 CC -1 FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS STALYU-LEWIS X.
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
 CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
 CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
 CC THE CELL SURFACE.
 CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1 SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS; BOVINE P-LECTIN LACKS
 CC THE HUMAN SUSHI-3, -4 AND -7 EQUIVALENTS.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L12041; AAA30743.1; -.
 DR HSSP: P16109; IFSB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR InterPro: IPR001304; Lectin_c.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; lectin_c; 1.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00032; CCP_6.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
 DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sush1; Repeat.
 FT SIGNAL 1 41
 FT CHAIN 42 646
 FT DOMAIN 42 587
 FT TRANSFEM 588 611
 FT DOMAIN 612 646
 FT DOMAIN 58 158
 FT DOMAIN 159 195
 FT DOMAIN 199 258
 FT DOMAIN 261 320
 FT DOMAIN 323 382
 FT DOMAIN 385 444

FT	DOMAIN	457	516	SUSHI 5.
FT	DOMAIN	519	578	SUSHI 6.
FT	DISULFID	60	158	BY SIMILARITY.
FT	DISULFID	131	150	BY SIMILARITY.
FT	DISULFID	163	174	BY SIMILARITY.
FT	DISULFID	168	183	BY SIMILARITY.
FT	DISULFID	185	194	BY SIMILARITY.
FT	DISULFID	200	244	BY SIMILARITY.
FT	DISULFID	230	257	BY SIMILARITY.
FT	DISULFID	262	306	BY SIMILARITY.
FT	DISULFID	292	319	BY SIMILARITY.
FT	DISULFID	324	368	BY SIMILARITY.
FT	DISULFID	354	381	BY SIMILARITY.
FT	DISULFID	386	430	BY SIMILARITY.
FT	DISULFID	416	443	BY SIMILARITY.
FT	DISULFID	458	502	BY SIMILARITY.
FT	DISULFID	488	515	BY SIMILARITY.
FT	DISULFID	520	564	BY SIMILARITY.
FT	DISULFID	550	577	BY SIMILARITY.
FT	CARBOHYD	48	48	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	54	54	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	80	80	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	180	180	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	212	212	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	219	219	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	336	336	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	481	481	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	532	532	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	539	539	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	557	557	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	634	637	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	646 AA;	71229 MW;	573912A4627A6ACA CRC64;

Query Match 39.68; Score 290; DB 1; Length 646;
 Best Local Similarity 39.08; Pred. No. 5; Se-21;
 Matches 48; Conservative 24; Mismatches 51; Indels 0; Gaps 0;

QY	1	CEPELAPLGLMDQTHPFNGFSFSCAFSCSEGTNTGTIEETGCPFGNWSPEPTQYV 60
DB	200	CGEPDLQGHVIMNCSHPLGNFSFSCAFSCSEGTNTGTIEETGCPFGNWSPEPTQYV 259
QY	61	IQCEPLAPDLGIMNCSHPLASFSFTACTGTCSEGTLEIGKKTKTCESSGIMNSPSPIC 120
DB	260	VQCPALKSPEGSMSCVQSAFAFQHQSSCSFSCBEGRALVGPVEVHTALGVMTATPPTVC 319
QY	121	QKL 123
DB	320	KAL 322

RESULT 18
 LEM2_RABIT
 ID LEM2_RABIT STANDARD; PRT; 551 AA.
 AC P27113;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
 DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
 DE (CD62E).
 GN SELE.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=92189729; PubMed=1372169;
 RA Lariagan J.D., Tsang T.C., Rumberger J.M., Burns D.K.;
 RT "Characterization of cDNA and genomic sequences encoding rabbit
 RT ELAM-1: conservation of structure and functional interactions with

RT leukocytes.";
 RL DNA Cell Biol. 11:149-162(1992).
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOLIPIDS).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- INDUCTION: BY CYTOKINES.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M91004; AAA31243.1; -;
 CC EMBL: M91005; AAA31244.1; -;
 CC HSSP: P16581; IKAJ.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR002396; Selectin.
 CC InterPro: IPR000436; Sush1_SCR_CCP.
 CC InterPro: IPR001304; lectin_c.
 CC Pfam: PF00008; EGF_1.
 CC Pfam: PF00059; lectin_c_1.
 CC Pfam: PF00084; sush1_5.
 CC PRINTS: PR00343; SELECTIN.
 CC SMART: SM00032; CCP_5.
 CC SMART: SM00034; CLECT_1.
 CC SMART: SM00181; EGF_1.
 CC PROSITE: PS00022; EGF_1_1.
 CC PROSITE: PS01186; EGF_2_1.
 CC PROSITE: PS00615; C-TYPE_LLECTIN_1.
 CC PROSITE: PS50041; C-TYPE_LLECTIN_2_1.
 CC Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sush1; Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 551 E-SELECTIN.
 FT DOMAIN 24 495 EXTRACELLULAR (POTENTIAL).
 FT TRANSMM 496 517 POTENTIAL.
 FT DOMAIN 518 551 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 40 140 C-TYPE LECTIN (SHORT FORM).
 FT DOMAIN 141 177 EGF-LIKE.
 FT DOMAIN 181 240 SUSHI 1.
 FT DOMAIN 243 302 SUSHI 2.
 FT DOMAIN 305 365 SUSHI 3.
 FT DOMAIN 368 428 SUSHI 4.
 FT DOMAIN 431 487 SUSHI 5.
 FT DISULFID 42 140 BY SIMILARITY.
 FT DISULFID 113 132 BY SIMILARITY.
 FT DISULFID 145 156 BY SIMILARITY.
 FT DISULFID 150 165 BY SIMILARITY.
 FT DISULFID 167 176 BY SIMILARITY.
 FT DISULFID 182 226 BY SIMILARITY.
 FT DISULFID 212 239 BY SIMILARITY.
 FT DISULFID 244 288 BY SIMILARITY.
 FT DISULFID 274 301 BY SIMILARITY.
 FT DISULFID 306 351 BY SIMILARITY.
 FT DISULFID 337 364 BY SIMILARITY.
 FT DISULFID 369 414 BY SIMILARITY.
 FT DISULFID 400 427 BY SIMILARITY.
 FT DISULFID 432 473 BY SIMILARITY.
 FT DISULFID 459 486 BY SIMILARITY.
 FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 308 308 T -> A (IN REF. 1; AAA31244).
 FT CONFLICT 328 328 T -> A (IN REF. 1; AAA31244).
 FT CONFLICT 491 491 A -> V (IN REF. 1; AAA31244).
 SQ SEQUENCE 551 AA; 60346 MW; 23BC8A883B23240E CRC64;
 Query Match 39.3%; Score 288; DB 1; Length 551;
 Best Local Similarity 40.5%; Pred. No. 7.3e-21;
 Matches 49; Conservative 20; Mismatches 52; Indels 0; Gaps 0;
 QY 1 CEPLAEPETGMDCTHPFNFSFSSQAFSCSGTNTLTGIEETGCPFGNMSPEPTCOV 60
 DB 182 CEAGQVOPQSGSLNCTHPNCFNFSYNSCSVSCERGIYLPSTETWTCTSSGEMSAHPATCY 241
 QY 61 IOCEPLAPDLGIMNCNHPPLASFSTACTFCSEGTLEIGRKRKTCSESGIMNSPSPIC 120
 DB 242 VECDFWGRKPRANGDVKCSFGSAPVMTCTCFPOEBEFTLLGARSLQCTSSGSDNEKPTC 301
 QY 121 Q 121
 DB 302 K 302
 RESULT 19
 LEM2_BOVIN
 ID LEM2_BOVIN STANDARD; PRT; 485 AA.
 AC P98107;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
 DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
 DE (CD62e).
 GN SELE.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adrenal gland;
 RX MEDLINE=93382537; PubMed=7690465;
 RA Nguyen M., Strubel N.A., Bischoff J.;
 RT "A role for sialyl Lewis-X/A glycoconjugates in capillary
 RT morphogenesis.";
 RT Nature 365:267-269(1993).
 RL -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOLIPIDS).
 CC -1- FUNCTION: INVOLVED IN CAPILLARY MORPHOGENESIS BY BINDING TO A
 CC BOVINE CAPILLARY ENDOTHELIAL (BCE) CELL SIALYL LEWIS X AND/OR
 CC SIALYL LEWIS A-CONTAINING LIGAND (PROBABLY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS; BOVINE E-LECTIN LACKS
 CC THE HUMAN SUSHI-4 AND 5 EQUIVALENTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; L12039; AAA02991.1; -.
DR HSSP; P16581; 1ESL.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 4.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 4.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE; PS00615; C-TYPE_LLECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 22
FT CHAIN 23 485
FT DOMAIN 23 430
FT TRANSMEM 431 453
FT DOMAIN 454 485
FT DOMAIN 39 139
FT DOMAIN 140 176
FT DOMAIN 180 238
FT DOMAIN 241 300
FT DOMAIN 303 363
FT DOMAIN 366 422
FT DOMAIN 41 139
FT DISULFID 112 131
FT DISULFID 144 155
FT DISULFID 149 164
FT DISULFID 166 175
FT DISULFID 181 224
FT DISULFID 210 237
FT DISULFID 242 286
FT DISULFID 272 299
FT DISULFID 304 349
FT DISULFID 335 362
FT DISULFID 367 408
FT DISULFID 394 421
FT CARBOHYD 61 61
FT CARBOHYD 79 79
FT CARBOHYD 88 88
FT CARBOHYD 161 161
FT CARBOHYD 203 203
FT CARBOHYD 265 265
FT CARBOHYD 312 312
FT CARBOHYD 316 316
FT CARBOHYD 379 379
FT CARBOHYD 401 401
SQ SEQUENCE 485 AA; 53200 MW; AE931C9B521E3904 CRC64;

Query Match 37.0%; Score 271.5; DB 1; Length 485;
Best Local Similarity 41.3%; Pred. No. 2,6e-19;
Matches 50; Conservative 20; Mismatches 50; Indels 1; Gaps 1;

OY 1 CEPLADELGTMDCTHPFGNFSFSSQCAFSCESTNLGTLEETTCGPGFMSSPEPTQY 60
DB 181 CPAQKHHEGHILVC-NELGKFTYNSSCISCAEGYLPSSTFATRCMSSGEMSTPLPCNV 239
OY 61 IOCEPLASAPDLGIMNCNHPPLASFPSFISACFIQSEGTGLGKKTKTIOESSGINSPEPIC 120
DB 240 VKCDALSLNDNGVNCSPNHSGLPWNTTCTFECEGKLTGPGHLOCTSSGIWDMKQPTC 299

OY 121 0 121
DB 300 K 300

RESULT 20
ID LEM2_PIG STANDARD; PRT; 484 AA.
AC P98110.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62e).
GN SELE.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=95071392; PubMed=7526854;
RA Rollins S.A., Evans M.J., Johnson K.K., Elliott E.A., Squinto S.P.,
RA Matis L.A., Rother R.P.;
RT "Molecular and functional analysis of porcine E-selectin reveals a
RT potential role in xenograft rejection."
RT Biochem. Biophys. Res. Commun. 204:763-771(1994).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=94271236; PubMed=7516159;
RA Tsang Y.T.M., Haskard D.O., Robinson M.K.;
RT "Cloning and expression kinetics of porcine vascular cell adhesion
RT molecule."
RT Biochem. Biophys. Res. Commun. 201:805-805(1994).
CC -!- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS STALKY-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC POLYACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN ACUTE CELLULAR ALLOGRAFT
CC REJECTION AND PROBABLY ALSO IN XENOGRAFT REJECTION.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS; PORCINE E-LECTIN LACKS
CC THE HUMAN SUSHI-1 AND -4 EQUIVALENTS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; L139076; AAA61545.1; -.
DR EMBL; U08350; AAA21541.1; -.
DR HSSP; P16581; 1ESL.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 4.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 4.
DR SMART; SM00034; CLECT; 1.

FT	SIGNAL	1	18	
FT	CHAIN	19	1231	COMPLEMENT FACTOR H.
FT	DOMAIN	20	81	SUSHI 1.
FT	DOMAIN	84	142	SUSHI 2.
FT	DOMAIN	145	206	SUSHI 3.
FT	DOMAIN	209	263	SUSHI 4.
FT	DOMAIN	266	321	SUSHI 5.
FT	DOMAIN	324	386	SUSHI 6.
FT	DOMAIN	388	443	SUSHI 7.
FT	DOMAIN	447	506	SUSHI 8.
FT	DOMAIN	508	565	SUSHI 9.
FT	DOMAIN	568	624	SUSHI 10.
FT	DOMAIN	629	685	SUSHI 11.
FT	DOMAIN	690	745	SUSHI 12.
FT	DOMAIN	752	804	SUSHI 13.
FT	DOMAIN	810	865	SUSHI 14.
FT	DOMAIN	869	927	SUSHI 15.
FT	DOMAIN	930	985	SUSHI 16.
FT	DOMAIN	988	1044	SUSHI 17.
FT	DOMAIN	1047	1103	SUSHI 18.
FT	DOMAIN	1108	1164	SUSHI 19.
FT	DOMAIN	1166	1229	SUSHI 20.
FT	DISULFID	21	66	BY SIMILARITY.
FT	DISULFID	52	80	BY SIMILARITY.
FT	DISULFID	85	129	BY SIMILARITY.
FT	DISULFID	114	141	BY SIMILARITY.
FT	DISULFID	146	192	BY SIMILARITY.
FT	DISULFID	178	205	BY SIMILARITY.
FT	DISULFID	210	251	BY SIMILARITY.
FT	DISULFID	237	262	BY SIMILARITY.
FT	DISULFID	267	309	BY SIMILARITY.
FT	DISULFID	294	320	BY SIMILARITY.
FT	DISULFID	325	374	BY SIMILARITY.
FT	DISULFID	357	385	BY SIMILARITY.
FT	DISULFID	389	431	BY SIMILARITY.
FT	DISULFID	416	442	BY SIMILARITY.
FT	DISULFID	448	494	BY SIMILARITY.
FT	DISULFID	477	505	BY SIMILARITY.
FT	DISULFID	509	553	BY SIMILARITY.
FT	DISULFID	536	564	BY SIMILARITY.
FT	DISULFID	569	611	BY SIMILARITY.
FT	DISULFID	597	623	BY SIMILARITY.
FT	DISULFID	630	673	BY SIMILARITY.
FT	DISULFID	659	684	BY SIMILARITY.
FT	DISULFID	691	733	BY SIMILARITY.
FT	DISULFID	719	744	BY SIMILARITY.
FT	DISULFID	753	792	BY SIMILARITY.
FT	DISULFID	781	803	BY SIMILARITY.
FT	DISULFID	811	853	BY SIMILARITY.
FT	DISULFID	839	864	BY SIMILARITY.
FT	DISULFID	870	915	BY SIMILARITY.
FT	DISULFID	901	926	BY SIMILARITY.
FT	DISULFID	931	973	BY SIMILARITY.
FT	DISULFID	959	984	BY SIMILARITY.
FT	DISULFID	989	1032	BY SIMILARITY.
FT	DISULFID	1018	1043	BY SIMILARITY.
FT	DISULFID	1048	1091	BY SIMILARITY.
FT	DISULFID	1077	1102	BY SIMILARITY.
FT	DISULFID	1109	1152	BY SIMILARITY.
FT	DISULFID	1138	1163	BY SIMILARITY.
FT	DISULFID	1167	1218	BY SIMILARITY.
FT	DISULFID	1201	1228	BY SIMILARITY.
FT	CARBOHYD	529	529	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	718	718	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	802	802	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	822	822	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	882	882	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	911	911	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	1029	1029	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	1095	1095	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	VARSPLIC	446	449	KTCS -> SETFL (IN ISOFORM 2).
FT	VARSPLIC	450	1231	H -> Y.
FT	VARIANT	402	402	

```

FT FT CONFLICT 21 21 /FTID=VAR_001979.
FF FF CONFLICT 30 30 C -> Q (IN REF. 3).
FF FF CONFLICT 34 34 T -> V (IN REF. 3).
FF FF CONFLICT 53 54 T -> Q (IN REF. 3).
FT FT STRAND 870 871 RP -> IL (IN REF. 2).
FT FT STRAND 876 876

Query Match 23.6%; Score 173; DB 1; Length 1231;
Best Local Similarity 28.8%; Pred. No. 2.7e-09;
Matches 36; Conservative 25; Mismatches 48; Indels 16; Gaps 5;

OY 14 CTH-----PRGNFS-----FSSQCAFSCSEGTMLTG-IEETTCGPGGNSSPEPTCOV 60
      | | | | | : : : | | : : : | | | : : : | | | : : : | | | : : : | | |
DB 85 CSHPDFTFGFTTGLGNGVNEFYGAAYVTCNMGYOLLEIGNVRECDTGD-WTNDIPICEV 143
OY 61 IOCEPLSPADLG-INMCSHPLASFISACFTICSETELLGKKKTCICSSGIGMSNSP 118
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 144 VACLEVTAPENKGVISAMEPDREYHFQAVREVCNKGKIEGDEMHCSDDGFWSKEKP 203
OY 119 ICQKL 123
      | : :
DB 204 KVEI 208

RESULT 23
CRL_HUMAN ~
ID CRL_HUMAN STANDARD: PRT: 2039 AA.
AC PI7927;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Complement receptor type 1 precursor (C3b/C4b receptor) (CD35
  antigen).
GN CRI OR C3BR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCRL_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89035992; PubMed=2972794;
RA Klickstein L.B., Bartow T.J., Miletic V., Rabson L.D., Smith J.A.,
RA Fearon D.T.;
RT "Identification of distinct C3b and C4b recognition sites in the
RT human C3b/C4b receptor (CRI, CD35) by deletion mutagenesis.";
RL J. Exp. Med. 168:1699-1717(1988).
RN [2]
RP SEQUENCE OF 503-2039 FROM N.A.
RX MEDLINE=87168191; PubMed=2951479;
RA Klickstein L.B., Wong W.W., Smith J.A., Wels J.H., Wilson J.G.,
RA Fearon D.T.;
RT "Human C3b/C4b receptor (CRI). Demonstration of long homologous
RT repeating domains that are composed of the short consensus repeats
RT characteristics of C3/C4 binding proteins.";
RL J. Exp. Med. 165:1095-1112(1987).
RN [3]
RP SEQUENCE OF 761-783, 831-845 AND 1179-1195 FROM N.A.
RX MEDLINE=86067975; PubMed=2933745;
RA Wong W.W., Klickstein L.B., Smith J.A., Wels J.H., Fearon D.T.;
RT "Identification of a partial cDNA clone for the human receptor for
RT complement fragments C3b/C4b ";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7711-7715(1985).
CC -!- FUNCTION: CRI, PRESENT ON ERYTHROCYTES, LEUKOCYTES, GLOMERULAR
CC PODOCYTES, AND SPLENIC FOLLICULAR DENDRITIC CELLS, MEDIATES THE
CC BINDING BY THESE CELLS OF PARTICLES AND IMMUNE COMPLEXES THAT HAVE
CC ACTIVATED COMPLEMENT.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- POLYMORPHISM: CRI IS RESPONSIBLE FOR THE KNOBS BLOOD GROUP SYSTEM.
CC -!- MISCELLANEOUS: SEVEN SHORT CONSENSUS REPEATS (SCR) CONSTITUTE A
CC LONG HOMOLOGOUS REPEAT (LHR). THE N-TERMINAL TWO SRCS OF LHR-A
CC CONTAINED A SITE DETERMINING CA SPECIFICITY, AND THE N-TERMINAL

```



```

FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 801 801 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1225 1225 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1234 AA; 139082 MW; C5AC02F341B957F7 CRC64;

Query Match 21.8%; Score 160; DB 1; Length 1234;
Best Local Similarity 28.0%; Pred. No. 5,1e-08;
Matches 35; Conservative 25; Mismatches 49; Indels 16; Gaps 5;

QY 14 CTH-----PFGN-----FSSQCAFSCSEGNLTG-IEETTCGPGNMSSPEPTQCV 60
DB 85 CGHPDGFPGFSGFRLVAGSOFEGAKVYTCDDGYOLGEIDYRECGADG-WINDIPICEV 143
QY 61 IOCEPLSAPDLG--IMNSHPLASFSTACTFICSECTELIGKTKTCSSGIMSNPSP 118
DB 144 VKCLPVTLENGRIYSGAETDQETFGQVYRFECSNGFKIEGHEIHCSENGIMSNKP 203
QY 119 ICOKL 123
DB 204 RCVEI 208

RESULT 25
DAFI_MOUSE STANDARD; PRT; 390 AA.
ID DAFI_MOUSE
AC 061475; 061397; P97732;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Complement decay-accelerating factor, GPI-anchored precursor
   (DAF-GPI).
GN DAF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=95403982; PubMed=7545711;
RA Spicer A.P., Seldin M.F., Gendler S.J.;
RT "Molecular cloning and chromosomal localization of the mouse decay-
   accelerating factor genes. Duplicated genes encode
   glycosylphosphatidylinositol-anchored and transmembrane forms.";
RL J. Immunol. 155:3079-3091(1995).
RN [2]
RP SEQUENCE OF 7-390 FROM N.A.
RC STRAIN=BALB/C; TISSUE=Spleen;
RX MEDLINE=96362213; PubMed=8671624;
RA Fukunaka Y., Yasui A., Okada H.;
RT "Molecular cloning of murine decay accelerating factor by
   immunoscreening.";
RL Int. Immunol. 8:379-385(1996).
CC -1- FUNCTION: PROTECTION OF CELLS FROM COMPLEMENT-MEDIATED DAMAGE (BY
   SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: BRAIN, SECRETORY EPITHELIA, SKELETAL MUSCLE,
   LIVER, TESTES, THYMUS, SPLEEN AND LYMPH NODE.
CC -1- DOMAIN: THE FIRST SUSHI DOMAIN (SCR1) IS NOT NECESSARY FOR
   FUNCTION. SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE
   ACTIVE SITE ON SCR3 (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
   (RCA) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L41366; AAB00091.1; -
DR EMBL; D63679; BAA09830.1; -
DR HSSP; P08603; IHCC.
DR MGD; MGI:104850; Dafi.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00064; sushi; 4.
DR SMART; SM00032; CCP; 4.
KW Complement pathway; Glycoprotein; Membrane; Repeat; GPI-anchor;
KW Signal; Sushi.
FT CHAIN 1 34
FT SIGNAL 35 361
FT PROPEP 362 390
FT DOMAIN 35 95
FT DOMAIN 97 159
FT DOMAIN 162 221
FT DOMAIN 224 285
FT DOMAIN 288 364
FT DISULFID 65 94
FT DISULFID 98 145
FT DISULFID 129 158
FT DISULFID 163 204
FT DISULFID 190 220
FT DISULFID 225 267
FT CARBOHYD 187 187
FT CARBOHYD 262 262
FT LIPID 361 361
FT CONFLICT 9 7
FT CONFLICT 83 9
FT CONFLICT 91 83
FT CONFLICT 135 135
FT CONFLICT 173 173
FT CONFLICT 180 180
SQ SEQUENCE 390 AA; 42618 MW; 4418721DFF478E7 CRC64;

Query Match 21.2%; Score 155.5; DB 1; Length 390;
Best Local Similarity 30.9%; Pred. No. 4,2e-08;
Matches 42; Conservative 16; Mismatches 63; Indels 15; Gaps 6;

QY 1 CEPLAPELLGTMCCTH---PFGNFSSSQCAFSCSGTNLTGIEETTCGPGN--WSSP 54
DB 158 CKKSCPNPKDLNGHINIPFG-TIFGSEINFCNPGYRLVGVSSTFCSTGNTYDWDDE 216
QY 55 EPTQVIOCEPLSAPDLGIMNSHPLASFSTACTFICSECTELIGKTKTCSSGIMSNPSP 110
DB 217 FVPCCTEIHCEPEPPKINGIMRGESD--SYTISQVYIYSCDKGFLVGNASTICTYKSDV 274
QY 111 GIWSNPSPICQKIDKS 126
DB 275 GQWSSPPRC--IEKS 288

RESULT 26
LFC_TACTR
ID LFC_TACTR STANDARD; PRT; 1019 AA.
AC P28175;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Limulus clotting factor C precursor (EC 3.4.21.84) (FC).
OS Limulus polyphemus (Japanese horseshoe crab).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC Limulidae; Tachyples.
OX NCBI_TaxID=6853;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

```

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%: Pred. No. 1, 1e-07; Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 0

Query	Db	Accession	Score	Length	Gap
29	FCSEGTNLGIEETTCGPGFNMWSSPEPTQYQICEPLADJGIMNCSHPLASFSTSA	866	21.2%	1019	
166	YSCSPGFKLGVARISCLPNCGOMSSFPPKC-IRECAKVSPEHGKVN--PSGNMIEGAT	22	34.7%	1019	
89	CTFICSGTELLGKKTKTIOESSGIMNSPICOIKL	123	33.0%	1019	
223	LRFSCDSPYLLIGETFLTCGNGMGSSGOIPCKKL	257	33.0%	1019	

Query Match

Best Local Similarity 34.7%:


```

CC -I- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
CC -I- SIMILARITY: TO C4B BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
CC -I- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-17 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31452; AAA36507.1; -.
DR EMBL; M62486; AAA36506.1; JOINED.
DR EMBL; M62475; AAA36506.1; JOINED.
DR EMBL; M62476; AAA36506.1; JOINED.
DR EMBL; M62477; AAA36506.1; JOINED.
DR EMBL; M62478; AAA36506.1; JOINED.
DR EMBL; M62479; AAA36506.1; JOINED.
DR EMBL; M62480; AAA36506.1; JOINED.
DR EMBL; M62481; AAA36506.1; JOINED.
DR EMBL; M62482; AAA36506.1; JOINED.
DR EMBL; M62484; AAA36506.1; JOINED.
DR EMBL; M62485; AAA36506.1; JOINED.
DR EMBL; M62485; AAA36506.1; JOINED.
DR EMBL; M62485; AAA36506.1; JOINED.
DR EMBL; X07853; CAA30701.1; -.
DR EMBL; X04284; CAA51244.1; -.
DR EMBL; X04296; CAA27839.1; -.
DR EMBL; X02865; CAA26617.1; -.
DR PIR; A33568; NBHUC4.
DR HSSP; P10998; IYVD.
DR MIM; 120830; -.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 8.
DR SMART; SM00032; CCP; 8.
KW Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal;
KW Polymorphism.
KW SIGNAL.
FT CHAIN 1 48
FT C4B-BINDING PROTEIN ALPHA CHAIN.
FT DOMAIN 49 597
FT SUSHI 1.
FT DOMAIN 112 171
FT SUSHI 2.
FT DOMAIN 174 235
FT SUSHI 3.
FT DOMAIN 238 295
FT SUSHI 4.
FT DOMAIN 298 361
FT SUSHI 5.
FT DOMAIN 364 423
FT SUSHI 6.
FT DOMAIN 425 481
FT SUSHI 7.
FT DOMAIN 483 539
FT SUSHI 8.
FT DISULFID 50 96
FT BY SIMILARITY.
FT DISULFID 81 108
FT BY SIMILARITY.
FT DISULFID 113 154
FT BY SIMILARITY.
FT DISULFID 140 170
FT BY SIMILARITY.
FT DISULFID 175 217
FT BY SIMILARITY.
FT DISULFID 203 234
FT BY SIMILARITY.
FT DISULFID 239 281
FT BY SIMILARITY.
FT DISULFID 267 294
FT BY SIMILARITY.
FT DISULFID 299 348
FT BY SIMILARITY.
FT DISULFID 332 360
FT BY SIMILARITY.
FT DISULFID 364 409
FT BY SIMILARITY.
FT DISULFID 7395 7409
FT BY SIMILARITY.
FT DISULFID 426 422
FT BY SIMILARITY.
FT DISULFID 454 468
FT BY SIMILARITY.
FT DISULFID 484 525
FT BY SIMILARITY.
FT DISULFID 511 538
FT BY SIMILARITY.
FT DISULFID 546 546
FT INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT DISULFID 558 558
FT INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT CARBOHYD 221 221
FT N-LINKED (GLCNAC. . .).
FT CARBOHYD 506 506
FT N-LINKED (GLCNAC. . .).
FT CARBOHYD 528 528
FT N-LINKED (GLCNAC. . .).
FT VARIANT 92 92
FT O -> T.
FT VARIANT 357 357
FT Y -> H.
FT VARIANT 473 473
FT /FTId=VAR_001978.
FT W -> L (IN DBSNP:1801341).

```



```

CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z50051; CA90391.1; -.
DR HSP; P10998; 1YVD.
DR Interpro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; sush1; 8.
DR SMART: SM00032; CCP: 8.
KM Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
FT SIGNAL 1 13 BY SIMILARITY.
FT CHAIN 14 558 C4B-BINDING PROTEIN ALPHA CHAIN.
FT DOMAIN 14 73 SUSH1 1.
FT DOMAIN 76 135 SUSH1 2.
FT DOMAIN 138 200 SUSH1 3.
FT DOMAIN 203 259 SUSH1 4.
FT DOMAIN 262 325 SUSH1 5.
FT DOMAIN 328 387 SUSH1 6.
FT DOMAIN 389 444 SUSH1 7.
FT DOMAIN 446 502 SUSH1 8.
FT DISULFID 15 60 BY SIMILARITY.
FT DISULFID 45 72 BY SIMILARITY.
FT DISULFID 77 118 BY SIMILARITY.
FT DISULFID 104 134 BY SIMILARITY.
FT DISULFID 139 182 BY SIMILARITY.
FT DISULFID 168 199 BY SIMILARITY.
FT DISULFID 204 246 BY SIMILARITY.
FT DISULFID 232 258 BY SIMILARITY.
FT DISULFID 263 312 BY SIMILARITY.
FT DISULFID 296 324 BY SIMILARITY.
FT DISULFID 328 351 BY SIMILARITY.
FT DISULFID 329 373 BY SIMILARITY.
FT DISULFID 363 386 BY SIMILARITY.
FT DISULFID 390 431 BY SIMILARITY.
FT DISULFID 417 443 BY SIMILARITY.
FT DISULFID 447 488 BY SIMILARITY.
FT DISULFID 474 501 BY SIMILARITY.
FT DISULFID 509 509 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT DISULFID 521 521 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 558 AA: 62266 MW: 5926067 EDIE5FF CRC64:
Query Match 19.9%; Score 146; DB 1; Length 558;
Best Local Similarity 28.6%; Pred. No. 5.2e-07;
Matches 42; Conservative 20; Mismatches 57; Indels 28; Gaps 7;
OY 1 CEPLEAPELG---TMDCTHPFG-----NFSFSCQAFSCSEGTNLGTIEETTC- 45
Db 60 CKPLCKWQININACVKRSCNPGDLONGKVEVKTDFLFGSIEFSCSEGVILIGSSSTYCE 119
OY 46 --GPRGNMSSPEPTCOVIOQCEPLSAPDLGIMNCSH---PLASEFTSACTFCSESTELI 100
Db 120 IQGKGVSSDPLPECVIACG--MPPD--ISNGKHNGREEFEFTYKSSVTVKCDPDTLL 175
OY 101 GKKTIC---ESSGIWNSPDPICOKL 123
Db 176 GNASTICVYVNTKGVWMSPTCERI 202

```

```

AC Q28065;
DR 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DI 16-OCT-2001 (rel. 40, Last annotation update)
DE C4B-binding protein alpha chain precursor (C4bp).
GN C4BPA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=liver;
RX MEDLINE=95015909; PubMed=7930621;
RA Hillarp A., Thern A., Dahlbaeck B.;
RT "Bovine C4b binding protein. Molecular cloning of the alpha- and
RT beta-chains provides structural background for lack of complex
RT formation with protein S."
RT J. Immunol. 153:4190-4199(1994).
CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
CC ACTIVATION. IT BINDS AS A COFACTOR TO C3B/C4B INACTIVATOR
CC (C3BINA), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4B. IT
CC ALSO ACCELERATES THE DEGRADATION OF THE C4B2A COMPLEX (C3
CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2A. ALPHA
CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.
CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG ALOPOLYPROTEIN R.
CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z31693; CA83498.1; -.
DR HSP; P10998; 1YVD.
DR Interpro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; sush1; 8.
DR SMART: SM00032; CCP: 8.
KM Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
FT SIGNAL 1 48 BY SIMILARITY.
FT CHAIN 49 610 C4B-BINDING PROTEIN ALPHA CHAIN.
FT DOMAIN 49 108 SUSH1 1.
FT DOMAIN 111 170 SUSH1 2.
FT DOMAIN 173 235 SUSH1 3.
FT DOMAIN 238 295 SUSH1 4.
FT DOMAIN 298 363 SUSH1 5.
FT DOMAIN 366 426 SUSH1 6.
FT DOMAIN 428 484 SUSH1 7.
FT DOMAIN 486 542 SUSH1 8.
FT DISULFID 50 95 BY SIMILARITY.
FT DISULFID 80 107 BY SIMILARITY.
FT DISULFID 112 153 BY SIMILARITY.
FT DISULFID 139 169 BY SIMILARITY.
FT DISULFID 174 217 BY SIMILARITY.
FT DISULFID 203 234 BY SIMILARITY.
FT DISULFID 239 281 BY SIMILARITY.
FT DISULFID 267 294 BY SIMILARITY.
FT DISULFID 299 350 BY SIMILARITY.
FT DISULFID 334 362 BY SIMILARITY.
FT DISULFID 366 390 BY SIMILARITY.
FT DISULFID 367 412 BY SIMILARITY.
FT DISULFID 402 425 BY SIMILARITY.
FT DISULFID 429 471 BY SIMILARITY.
FT DISULFID 457 483 BY SIMILARITY.
FT DISULFID 487 528 BY SIMILARITY.
FT DISULFID 514 541 BY SIMILARITY.
FT DISULFID 549 549 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).

```

```

FT DISULFID 561 561 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 610 AA: 68886 MW: D806B270E8A06B58 CRC64;

Query Match 19.8%; Score 145; DB 1; Length 610;
Best Local Similarity 30.7%; Pred. No. 7.1e-07;
Matches 35; Conservative 18; Mismatches 47; Indels 14; Gaps 4;

OY 20 NFSESCAFSCSEGTNLGIEETTCGPPGN---WSSEPTCQVIGCEPLAPDGIIMNC 76
DB 128 DYFSGSEIEFSCSEGYVIGSANSYCOLDKGVWSDPLPQCIIAKCEPPPT---ISNG 183
OY 77 SHPLAS---FSFTSACFTICSGTELGKKTIC---ESSGIWNSPPIQOKL 123
DB 184 RHNGDEDFPYTGSSVITSCDRFSMLGKASISCHVENKKTIGWSPSPCKKV 237

RESULT 32
C4BP_MOUSE STANDARD: PRT: 469 AA.
ID C4BP_MOUSE AC P06607;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C4b-binding protein precursor (C4bp).
GN C4BPA OR C4BP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88024997; PubMed=3663616;
RA Kristensen T., Ogata R.T., Chung L.P., Reid K.B.M., Tack B.F.;
RT "cDNA structure of murine C4b-binding protein, a regulatory component
of the serum complement system."
RL Biochemistry 26:4668-4674(1987).
CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
(C3bINA), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4b. IT
ALSO ACCELERATES THE DEGRADATION OF THE C4bC2a COMPLEX (C3
CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. ALPHA
CHAIN BINDS C4b. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.
CC -1- SUBUNIT: HOMOPENTAMER, NOT COVALENTLY LINKED. MOUSE LACKS THE
BETA CHAIN OF C4BP.
CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-44 IS THE INITIATOR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17122; AAA37312.1; ALT_INIT.
DR PIR: A27117; NEMSC4.
DR HSSP: P10996; IVD.
DR MGD: MGI:88229; C4bp.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; sush1: 6.
DR SMART: SM00032; CCP; 6.
KW Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
FT SIGNAL 1 56
FT CHAIN 57 469 C4B-BINDING PROTEIN.
FT DOMAIN 57 116 SUSHI 1.
FT DOMAIN 119 177 SUSHI 2.

```

```

FT DOMAIN 180 241 SUSHI 3.
FT DOMAIN 244 300 SUSHI 4.
FT DOMAIN 302 356 SUSHI 5.
FT DOMAIN 358 414 SUSHI 6.
FT DISULFID 58 103 BY SIMILARITY.
FT DISULFID 88 115 BY SIMILARITY.
FT DISULFID 120 160 BY SIMILARITY.
FT DISULFID 146 176 BY SIMILARITY.
FT DISULFID 181 223 BY SIMILARITY.
FT DISULFID 209 240 BY SIMILARITY.
FT DISULFID 245 287 BY SIMILARITY.
FT DISULFID 273 299 BY SIMILARITY.
FT DISULFID 303 343 BY SIMILARITY.
FT DISULFID 329 355 BY SIMILARITY.
FT DISULFID 359 400 BY SIMILARITY.
FT DISULFID 386 413 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 428 428 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 469 AA: 51551 MW: 41E137CB8D8C6321 CRC64;

Query Match 19.7%; Score 144.5; DB 1; Length 469;
Best Local Similarity 29.4%; Pred. No. 6.1e-07;
Matches 32; Conservative 19; Mismatches 49; Indels 9; Gaps 3;

OY 22 SFSSQCAFSCSEGTNLGIEETTCGPPGN---WSSEPTCQVIGCEPLAPDGIIMNC 78
DB 137 TFGSQIEFSCSEGYVIGSSTSCSEYVKGAVMNSNPPELVYKCGP--PPDISNKHSG 194
OY 79 PLASFSTSACFTICSGTELGKKTIC---ESSGIWNSPPIQOKL 123
DB 195 TEDFPYNHGISTYCDPGRFVGSFPGCTVNVKTPVWSSPPTCEKI 243

RESULT 33
C4BP_RAT STANDARD: PRT: 258 AA.
ID C4BP_RAT AC Q63515;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE C4b-binding protein beta chain precursor.
GN C4BPB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE=97160682; PubMed=9013975;
RA Hillarp A., Wiklund H., Thern A., Dahlback B.;
RT "Molecular cloning of rat C4b binding protein alpha- and beta-chains:
structural and functional relationships among human, bovine, rabbit,
mouse, and rat proteins."
RL J. Immunol. 158:1315-1323(1997).
CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
(C3bINA), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4b. IT
ALSO ACCELERATES THE DEGRADATION OF THE C4bC2a COMPLEX (C3
CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. IT
INTERACTS ALSO WITH ANTICOAGULANT PROTEIN S AND WITH SERUM AMYLOID
P COMPONENT.
CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: TO C4BP ALPHA CHAIN AND TO PIG APOLIPOPROTEIN R.
CC -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) DOMAINS.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration

```



```

CC CC DISOCIATION OF C3 CONVERTASE.
CC -I- SUBCELLULAR LOCATION: CLASS GPI: ATTACHED TO THE MEMBRANE BY A
CC GPI-ANCHOR.
CC -I- ALTERNATIVE PRODUCTS: THERE ARE AT LEAST SIX FORMS OF DAF
CC TRANSCRIPTS (DISTRIBUTED IN FOUR CLASSES: GPI, TCL, TCS AND SEC)
CC GENERATED BY ALTERNATIVE SPLICING. ONE GPI-ANCHORED FORM (GDAB-
CC GPI), TWO FORMS WITH LONGER CYTOPLASMIC REGION (GDAB-TCL AND
CC GDAB-C-TCL), TWO FORMS WITH SHORTER CYTOPLASMIC REGION (GDA-TCS AND
CC GDAB-TCS) AND ONE SECRETED FORM (GDAB-SEC).
CC -I- TISSUE SPECIFICITY: ALL THE ISOFORMS ARE WIDELY EXPRESSED. GPI AND
CC TCS ARE THE MAJOR FORMS, WHEREAS SEC IS MINOR AND TCL IS ONLY
CC PRESENT IN TRACE LEVELS.
CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCS) DOMAINS.
CC -I- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
CC (RCA) FAMILY.
CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: D55667; BAA09514.1; JOINED.
CC DR EMBL: D55656; BAA09514.1; JOINED.
CC DR EMBL: D55657; BAA09514.1; JOINED.
CC DR EMBL: D55658; BAA09514.1; JOINED.
CC DR EMBL: D55659; BAA09514.1; JOINED.
CC DR EMBL: D55660; BAA09514.1; JOINED.
CC DR EMBL: D55661; BAA09514.1; JOINED.

```

DR	EMBL	D55664	BAA09514.1	JOINED
DR	EMBL	D55667	BAA09515.1	-
DR	EMBL	D55666	BAA09515.1	JOINED
DR	EMBL	D55657	BAA09515.1	JOINED
DR	EMBL	D55658	BAA09515.1	JOINED
DR	EMBL	D55659	BAA09515.1	JOINED
DR	EMBL	D55660	BAA09515.1	JOINED
DR	EMBL	D55661	BAA09515.1	JOINED
DR	EMBL	D55662	BAA09515.1	JOINED
DR	EMBL	D55663	BAA09515.1	JOINED
DR	EMBL	D55664	BAA09515.1	JOINED
DR	EMBL	D55665	BAA09515.1	JOINED
DR	EMBL	D55667	BAA09516.1	-
DR	EMBL	D55656	BAA09516.1	JOINED
DR	EMBL	D55657	BAA09516.1	JOINED
DR	EMBL	D55658	BAA09516.1	JOINED
DR	EMBL	D55659	BAA09516.1	JOINED
DR	EMBL	D55660	BAA09516.1	JOINED
DR	EMBL	D55661	BAA09516.1	JOINED
DR	EMBL	D55662	BAA09516.1	JOINED
DR	EMBL	D55663	BAA09516.1	JOINED
DR	EMBL	D55664	BAA09516.1	JOINED
DR	EMBL	D55665	BAA09516.1	JOINED
DR	EMBL	D55667	BAA09517.1	-
DR	EMBL	D55656	BAA09517.1	JOINED
DR	EMBL	D55657	BAA09517.1	JOINED
DR	EMBL	D55658	BAA09517.1	JOINED
DR	EMBL	D55659	BAA09517.1	JOINED
DR	EMBL	D55660	BAA09517.1	JOINED
DR	EMBL	D55661	BAA09517.1	JOINED
DR	EMBL	D55662	BAA09517.1	JOINED
DR	EMBL	D55663	BAA09517.1	JOINED
DR	EMBL	D55664	BAA09517.1	JOINED
DR	EMBL	D55665	BAA09517.1	JOINED
DR	EMBL	D55667	BAA09518.1	-
DR	EMBL	D55656	BAA09518.1	JOINED
DR	EMBL	D55657	BAA09518.1	JOINED
DR	EMBL	D55658	BAA09518.1	JOINED
DR	EMBL	D55659	BAA09518.1	JOINED

DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR Pfam: PF00084; sush1. 4.
 DR SMART: SM00032; CCP; 4.
 KW Heparin-binding; Glycoprotein; Plasma; Repeat; Sush1; Signal.
 FT SIGNAL 1
 FT CHAIN 1 19
 FT DOMAIN 20 345
 FT DOMAIN 22 80
 FT DOMAIN 83 138
 FT DOMAIN 141 201
 FT DOMAIN 204 261
 FT DOMAIN 263 345
 FT DISULFID 23 66
 FT DISULFID 51 79
 FT DISULFID 84 124
 FT DISULFID 110 137
 FT DISULFID 142 188
 FT DISULFID 174 200
 FT DISULFID 205 248
 FT DISULFID 234 260
 FT DISULFID 264 315
 FT DISULFID 300 325
 FT DISULFID 307 345
 FT CARBOHYD 92 92
 FT CARBOHYD 162 162
 FT CARBOHYD 183 183
 FT CARBOHYD 193 193
 FT CARBOHYD 253 253
 FT CONFLICT 101 101
 FT CONFLICT 108 108
 FT CONFLICT 177 177
 FT CONFLICT 191 191
 FT CONFLICT 194 194
 FT CONFLICT 259 259
 FT CONFLICT 302 302
 FT CONFLICT 305 305
 FT CONFLICT 329 329
 SQ SEQUENCE 345 AA; 38252 MW; E117DAB609461C33 CRC64;
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 E -> G (IN REF. 1).
 F -> S (IN REF. 1).
 H -> R (IN REF. 1).
 H -> N (IN REF. 3).
 W -> C (IN REF. 1).
 S -> N (IN REF. 1).
 H -> N (IN REF. 1).
 K -> R (IN REF. 1).
 H -> R (IN REF. 1).
 Query Match 19.3%; Score 141.5; DB 1; Length 345;
 Best Local Similarity 31.5%; Pred. No. 8.7e-07;
 Matches 39; Conservative 15; Mismatches 63; Indels 7; Gaps 6;
 QY 1 CEPLAEPLGTMDCTHPF-GNFSF-SSQCAFSCSEGTNLGTIEETTCGPFGNMSPEPIC 58
 DB 142 CPEPPPTPKASLSVYKFLGNSFYGSKAVFKCLPHHAFGNDVTCTEHNWQ-LPEC 200
 QY 59 QYIQCEPLSAPDLGTMNCSHPL-ASFSTSACTFICSEGTGLGKKTKICSSGINSNPS 117
 DB 201 REVRCPFPSPRDPNGFVN--HPANPVLYKDTATFGCHETYSLDGPREVECSKFGNMS-AQ 257
 QY 118 PTQC 121
 DB 258 PSCK 261
 RESULT 36
 CR2_HUMAN
 ID CR2_HUMAN STANDARD; PRT; 1033 AA.
 AC P20023;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Complement receptor type 2 precursor (Cr2) (Complement C3d receptor)
 DE (Epstein-Barr virus receptor) (EBV receptor) (CD21 antigen).
 GN CR2 OR C3DR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP MEDLINE=89123277; PubMed=2563370;
 RX

RA Fujisaku A., Harley J.B., Frank M.B., Gruner B.A., Frazier B.,
 RA Holers V.M.;
 RT "Genomic organization and polymorphisms of the human C3d/Epstein-Barr
 RT virus receptor.";
 RL J. Biol. Chem. 264:2118-2125(1989).
 RN [2]
 RP SEQUENCE OF 226-233; 256-267; 332-341; 667-677 AND 898-908.
 RX MEDLINE=86287311; PubMed=3016712;
 RA Wels J.J., Fearon D.T., Klickstein L.B., Wong W.W., Richards S.A.,
 RA de Bruyn Kops A., Smith J.A., Wels J.H.;
 RT "Identification of a partial cDNA clone for the C3d/Epstein-Barr
 RT virus receptor of human B lymphocytes: homology with the receptor for
 RT fragments C3b and C4b of the third and fourth components of
 RT complement.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5639-5643(1986).
 RN [3]
 RP SEQUENCE OF 492-556 FROM N.A. (SHORT ISOFORM).
 RX MEDLINE=93294286; PubMed=8390533;
 RA Sinha S.K., Todd S.C., Hedrick J.A., Spelsberg C.L., Lambiris J.D.,
 RA Tsoukas C.D.;
 RT "Characterization of the EBV/C3d receptor on the human Jurkat T cell
 RT line: evidence for a novel transcript.";
 RL J. Immunol. 150:5311-5320(1993).
 CC - FUNCTION: RECEPTOR FOR COMPLEMENT C3D AND FOR THE EPSTEIN-BARR
 CC VIRUS ON HUMAN B-CELLS AND T-CELLS. PARTICIPATES IN B LYMPHOCYTES
 CC ACTIVATION.
 CC - SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC - TISSUE SPECIFICITY: MATURE B LYMPHOCYTES, T LYMPHOCYTES AND
 CC FOLLICULAR DENDRITIC CELLS OF THE SPLEEN.
 CC - SIMILARITY: TO MOUSE CR2. CD21 IS A MEMBER OF THE REGULATORS OF
 CC COMPLEMENT ACTIVATION (RCA) FAMILY.
 CC - SIMILARITY: CONTAINS 15 SUSHI (SCR) DOMAINS.
 CC - DATABASE: NAME=PROW; NOTE=CD guide CD21 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd21.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: M26004; AAA35786.1; -;
 DR EMBL: M26016; AAB04638.1; -;
 DR EMBL: M24007; AAB04638.1; JOINED.
 DR EMBL: M24008; AAB04638.1; JOINED.
 DR EMBL: M24009; AAB04638.1; JOINED.
 DR EMBL: M24010; AAB04638.1; JOINED.
 DR EMBL: M24011; AAB04638.1; JOINED.
 DR EMBL: M26009; AAB04638.1; JOINED.
 DR EMBL: M26010; AAB04638.1; JOINED.
 DR EMBL: M26011; AAB04638.1; JOINED.
 DR EMBL: M26012; AAB04638.1; JOINED.
 DR EMBL: M26013; AAB04638.1; JOINED.
 DR EMBL: M26014; AAB04638.1; JOINED.
 DR EMBL: M26015; AAB04638.1; JOINED.
 DR EMBL: M26016; AAB04638.1; JOINED.
 DR EMBL: M26017; AAB04638.1; JOINED.
 DR EMBL: M26018; AAB04638.1; JOINED.
 DR EMBL: M26019; AAB04638.1; JOINED.
 DR EMBL: M26020; AAB04638.1; JOINED.
 DR EMBL: M26021; AAB04638.1; JOINED.
 DR EMBL: M26022; AAB04638.1; JOINED.
 DR EMBL: M26023; AAB04638.1; JOINED.
 DR EMBL: M26024; AAB04638.1; JOINED.
 DR EMBL: M26025; AAB04638.1; JOINED.
 DR EMBL: M26026; AAB04638.1; JOINED.
 DR EMBL: M26027; AAB04638.1; JOINED.
 DR EMBL: M26028; AAB04638.1; JOINED.
 DR EMBL: M26029; AAB04638.1; JOINED.
 DR EMBL: M26030; AAB04638.1; JOINED.
 DR EMBL: M26031; AAB04638.1; JOINED.
 DR EMBL: M26032; AAB04638.1; JOINED.
 DR EMBL: M26033; AAB04638.1; JOINED.
 DR EMBL: M26034; AAB04638.1; JOINED.
 DR EMBL: M26035; AAB04638.1; JOINED.
 DR EMBL: M26036; AAB04638.1; JOINED.
 DR EMBL: M26037; AAB04638.1; JOINED.
 DR EMBL: M26038; AAB04638.1; JOINED.
 DR EMBL: M26039; AAB04638.1; JOINED.
 DR EMBL: M26040; AAB04638.1; JOINED.
 DR EMBL: M26041; AAB04638.1; JOINED.
 DR EMBL: M26042; AAB04638.1; JOINED.
 DR EMBL: M26043; AAB04638.1; JOINED.
 DR EMBL: M26044; AAB04638.1; JOINED.
 DR EMBL: M26045; AAB04638.1; JOINED.
 DR EMBL: M26046; AAB04638.1; JOINED.
 DR EMBL: M26047; AAB04638.1; JOINED.
 DR EMBL: M26048; AAB04638.1; JOINED.
 DR EMBL: M26049; AAB04638.1; JOINED.
 DR EMBL: M26050; AAB04638.1; JOINED.
 DR EMBL: M26051; AAB04638.1; JOINED.
 DR EMBL: M26052; AAB04638.1; JOINED.
 DR EMBL: M26053; AAB04638.1; JOINED.
 DR EMBL: M26054; AAB04638.1; JOINED.
 DR EMBL: M26055; AAB04638.1; JOINED.
 DR EMBL: M26056; AAB04638.1; JOINED.
 DR EMBL: M26057; AAB04638.1; JOINED.
 DR EMBL: M26058; AAB04638.1; JOINED.
 DR EMBL: M26059; AAB04638.1; JOINED.
 DR EMBL: M26060; AAB04638.1; JOINED.
 DR EMBL: M26061; AAB04638.1; JOINED.
 DR EMBL: M26062; AAB04638.1; JOINED.
 DR EMBL: M26063; AAB04638.1; JOINED.
 DR EMBL: M26064; AAB04638.1; JOINED.
 DR EMBL: M26065; AAB04638.1; JOINED.
 DR EMBL: M26066; AAB04638.1; JOINED.
 DR EMBL: M26067; AAB04638.1; JOINED.
 DR EMBL: M26068; AAB04638.1; JOINED.
 DR EMBL: M26069; AAB04638.1; JOINED.
 DR EMBL: M26070; AAB04638.1; JOINED.
 DR EMBL: M26071; AAB04638.1; JOINED.
 DR EMBL: M26072; AAB04638.1; JOINED.
 DR EMBL: M26073; AAB04638.1; JOINED.
 DR EMBL: M26074; AAB04638.1; JOINED.
 DR EMBL: M26075; AAB04638.1; JOINED.
 DR EMBL: M26076; AAB04638.1; JOINED.
 DR EMBL: M26077; AAB04638.1; JOINED.
 DR EMBL: M26078; AAB04638.1; JOINED.
 DR EMBL: M26079; AAB04638.1; JOINED.
 DR EMBL: M26080; AAB04638.1; JOINED.
 DR EMBL: M26081; AAB04638.1; JOINED.
 DR EMBL: M26082; AAB04638.1; JOINED.
 DR EMBL: M26083; AAB04638.1; JOINED.
 DR EMBL: M26084; AAB04638.1; JOINED.
 DR EMBL: M26085; AAB04638.1; JOINED.
 DR EMBL: M26086; AAB04638.1; JOINED.
 DR EMBL: M26087; AAB04638.1; JOINED.
 DR EMBL: M26088; AAB04638.1; JOINED.
 DR EMBL: M26089; AAB04638.1; JOINED.
 DR EMBL: M26090; AAB04638.1; JOINED.
 DR EMBL: M26091; AAB04638.1; JOINED.
 DR EMBL: M26092; AAB04638.1; JOINED.
 DR EMBL: M26093; AAB04638.1; JOINED.
 DR EMBL: M26094; AAB04638.1; JOINED.
 DR EMBL: M26095; AAB04638.1; JOINED.
 DR EMBL: M26096; AAB04638.1; JOINED.
 DR EMBL: M26097; AAB04638.1; JOINED.
 DR EMBL: M26098; AAB04638.1; JOINED.
 DR EMBL: M26099; AAB04638.1; JOINED.
 DR EMBL: M26100; AAB04638.1; JOINED.
 DR EMBL: M26101; AAB04638.1; JOINED.
 DR EMBL: M26102; AAB04638.1; JOINED.
 DR EMBL: M26103; AAB04638.1; JOINED.
 DR EMBL: M26104; AAB04638.1; JOINED.
 DR EMBL: M26105; AAB04638.1; JOINED.
 DR EMBL: M26106; AAB04638.1; JOINED.
 DR EMBL: M26107; AAB04638.1; JOINED.
 DR EMBL: M26108; AAB04638.1; JOINED.
 DR EMBL: M26109; AAB04638.1; JOINED.
 DR EMBL: M26110; AAB04638.1; JOINED.
 DR EMBL: M26111; AAB04638.1; JOINED.
 DR EMBL: M26112; AAB04638.1; JOINED.
 DR EMBL: M26113; AAB04638.1; JOINED.
 DR EMBL: M26114; AAB04638.1; JOINED.
 DR EMBL: M26115; AAB04638.1; JOINED.
 DR EMBL: M26116; AAB04638.1; JOINED.
 DR EMBL: M26117; AAB04638.1; JOINED.
 DR EMBL: M26118; AAB04638.1; JOINED.
 DR EMBL: M26119; AAB04638.1; JOINED.
 DR EMBL: M26120; AAB04638.1; JOINED.
 DR EMBL: M26121; AAB04638.1; JOINED.
 DR EMBL: M26122; AAB04638.1; JOINED.
 DR EMBL: M26123; AAB04638.1; JOINED.
 DR EMBL: M26124; AAB04638.1; JOINED.
 DR EMBL: M26125; AAB04638.1; JOINED.
 DR EMBL: M26126; AAB04638.1; JOINED.
 DR EMBL: M26127; AAB04638.1; JOINED.
 DR EMBL: M26128; AAB04638.1; JOINED.
 DR EMBL: M26129; AAB04638.1; JOINED.
 DR EMBL: M26130; AAB04638.1; JOINED.
 DR EMBL: M26131; AAB04638.1; JOINED.
 DR EMBL: M26132; AAB04638.1; JOINED.
 DR EMBL: M26133; AAB04638.1; JOINED.
 DR EMBL: M26134; AAB04638.1; JOINED.
 DR EMBL: M26135; AAB04638.1; JOINED.
 DR EMBL: M26136; AAB04638.1; JOINED.
 DR EMBL: M26137; AAB04638.1; JOINED.
 DR EMBL: M26138; AAB04638.1; JOINED.
 DR EMBL: M26139; AAB04638.1; JOINED.
 DR EMBL: M26140; AAB04638.1; JOINED.
 DR EMBL: M26141; AAB04638.1; JOINED.
 DR EMBL: M26142; AAB04638.1; JOINED.
 DR EMBL: M26143; AAB04638.1; JOINED.
 DR EMBL: M26144; AAB04638.1; JOINED.
 DR EMBL: M26145; AAB04638.1; JOINED.
 DR EMBL: M26146; AAB04638.1; JOINED.
 DR EMBL: M26147; AAB04638.1; JOINED.
 DR EMBL: M26148; AAB04638.1; JOINED.
 DR EMBL: M26149; AAB04638.1; JOINED.
 DR EMBL: M26150; AAB04638.1; JOINED.
 DR EMBL: M26151; AAB04638.1; JOINED.
 DR EMBL: M26152; AAB04638.1; JOINED.
 DR EMBL: M26153; AAB04638.1; JOINED.
 DR EMBL: M26154; AAB04638.1; JOINED.
 DR EMBL: M26155; AAB04638.1; JOINED.
 DR EMBL: M26156; AAB04638.1; JOINED.
 DR EMBL: M26157; AAB04638.1; JOINED.
 DR EMBL: M26158; AAB04638.1; JOINED.
 DR EMBL: M26159; AAB04638.1; JOINED.
 DR EMBL: M26160; AAB04638.1; JOINED.
 DR EMBL: M26161; AAB04638.1; JOINED.
 DR EMBL: M26162; AAB04638.1; JOINED.
 DR EMBL: M26163; AAB04638.1; JOINED.
 DR EMBL: M26164; AAB04638.1; JOINED.
 DR EMBL: M26165; AAB04638.1; JOINED.
 DR EMBL: M26166; AAB04638.1; JOINED.
 DR EMBL: M26167; AAB04638.1; JOINED.
 DR EMBL: M26168; AAB04638.1; JOINED.
 DR EMBL: M26169; AAB04638.1; JOINED.
 DR EMBL: M26170; AAB04638.1; JOINED.
 DR EMBL: M26171; AAB04638.1; JOINED.
 DR EMBL: M26172; AAB04638.1; JOINED.
 DR EMBL: M26173; AAB04638.1; JOINED.
 DR EMBL: M26174; AAB04638.1; JOINED.
 DR EMBL: M26175; AAB04638.1; JOINED.
 DR EMBL: M26176; AAB04638.1; JOINED.
 DR EMBL: M26177; AAB04638.1; JOINED.
 DR EMBL: M26178; AAB04638.1; JOINED.
 DR EMBL: M26179; AAB04638.1; JOINED.
 DR EMBL: M26180; AAB04638.1; JOINED.
 DR EMBL: M26181; AAB04638.1; JOINED.
 DR EMBL: M26182; AAB04638.1; JOINED.
 DR EMBL: M26183; AAB04638.1; JOINED.
 DR EMBL: M26184; AAB04638.1; JOINED.
 DR EMBL: M26185; AAB04638.1; JOINED.
 DR EMBL: M26186; AAB04638.1; JOINED.
 DR EMBL: M26187; AAB04638.1; JOINED.
 DR EMBL: M26188; AAB04638.1; JOINED.
 DR EMBL: M26189; AAB04638.1; JOINED.
 DR EMBL: M26190; AAB04638.1; JOINED.
 DR EMBL: M26191; AAB04638.1; JOINED.
 DR EMBL: M26192; AAB04638.1; JOINED.
 DR EMBL: M26193; AAB04638.1; JOINED.
 DR EMBL: M26194; AAB04638.1; JOINED.
 DR EMBL: M26195; AAB04638.1; JOINED.
 DR EMBL: M26196; AAB04638.1; JOINED.
 DR EMBL: M26197; AAB04638.1; JOINED.
 DR EMBL: M26198; AAB04638.1; JOINED.
 DR EMBL: M26199; AAB04638.1; JOINED.
 DR EMBL: M26200; AAB04638.1; JOINED.
 DR EMBL: M26201; AAB04638.1; JOINED.
 DR EMBL: M26202; AAB04638.1; JOINED.
 DR EMBL: M26203; AAB04638.1; JOINED.
 DR EMBL: M26204; AAB04638.1; JOINED.
 DR EMBL: M26205; AAB04638.1; JOINED.
 DR EMBL: M26206; AAB04638.1; JOINED.
 DR EMBL: M26207; AAB04638.1; JOINED.
 DR EMBL: M26208; AAB04638.1; JOINED.
 DR EMBL: M26209; AAB04638.1; JOINED.
 DR EMBL: M26210; AAB04638.1; JOINED.
 DR EMBL: M26211; AAB04638.1; JOINED.
 DR EMBL: M26212; AAB04638.1; JOINED.
 DR EMBL: M26213; AAB04638.1; JOINED.
 DR EMBL: M26214; AAB04638.1; JOINED.
 DR EMBL: M26215; AAB04638.1; JOINED.
 DR EMBL: M26216; AAB04638.1; JOINED.
 DR EMBL: M26217; AAB04638.1; JOINED.
 DR EMBL: M26218; AAB04638.1; JOINED.
 DR EMBL: M26219; AAB04638.1; JOINED.
 DR EMBL: M26220; AAB04638.1; JOINED.
 DR EMBL: M26221; AAB04638.1; JOINED.
 DR EMBL: M26222; AAB04638.1; JOINED.
 DR EMBL: M26223; AAB04638.1; JOINED.
 DR EMBL: M26224; AAB04638.1; JOINED.
 DR EMBL: M26225; AAB04638.1; JOINED.
 DR EMBL: M26226; AAB04638.1; JOINED.
 DR EMBL: M26227; AAB04638.1; JOINED.
 DR EMBL: M26228; AAB04638.1; JOINED.
 DR EMBL: M26229; AAB04638.1; JOINED.
 DR EMBL: M26230; AAB04638.1; JOINED.
 DR EMBL: M26231; AAB04638.1; JOINED.
 DR EMBL: M26232; AAB04638.1; JOINED.
 DR EMBL: M26233; AAB04638.1; JOINED.
 DR EMBL: M26234; AAB04638.1; JOINED.
 DR EMBL: M26235; AAB04638.1; JOINED.
 DR EMBL: M26236; AAB04638.1; JOINED.
 DR EMBL: M26237; AAB04638.1; JOINED.
 DR EMBL: M26238; AAB04638.1; JOINED.
 DR EMBL: M26239; AAB04638.1; JOINED.
 DR EMBL: M26240; AAB04638.1; JOINED.
 DR EMBL: M26241; AAB04638.1; JOINED.
 DR EMBL: M26242; AAB04638.1; JOINED.
 DR EMBL: M26243; AAB04638.1; JOINED.
 DR EMBL: M26244; AAB04638.1; JOINED.
 DR EMBL: M26245; AAB04638.1; JOINED.
 DR EMBL: M26246; AAB04638.1; JOINED.
 DR EMBL: M26247; AAB04638.1; JOINED.
 DR EMBL: M26248; AAB04638.1; JOINED.
 DR EMBL: M26249; AAB04638.1; JOINED.
 DR EMBL: M26250; AAB04638.1; JOINED.
 DR EMBL: M26251; AAB04638.1; JOINED.
 DR EMBL: M26252; AAB04638.1; JOINED.
 DR EMBL: M26253; AAB04638.1; JOINED.
 DR EMBL: M26254; AAB04638.1; JOINED.
 DR EMBL: M26255; AAB04638.1; JOINED.
 DR EMBL: M26256; AAB04638.1; JOINED.
 DR EMBL: M26257; AAB04638.1; JOINED.
 DR EMBL: M26258; AAB04638.1; JOINED.
 DR EMBL: M26259; AAB04638.1; JOINED.
 DR EMBL: M26260; AAB04638.1; JOINED.
 DR EMBL: M26261; AAB04638.1; JOINED.
 DR EMBL: M26262; AAB04638.1; JOINED.
 DR EMBL: M26263; AAB04638.1; JOINED.
 DR EMBL: M26264; AAB04638.1; JOINED.
 DR EMBL: M26265; AAB04638.1; JOINED.
 DR EMBL: M26266; AAB04638.1; JOINED.
 DR EMBL: M26267; AAB04638.1; JOINED.
 DR EMBL: M26268; AAB04638.1; JOINED.
 DR EMBL: M26269; AAB04638.1; JOINED.
 DR EMBL: M26270; AAB04638.1; JOINED.
 DR EMBL: M26271; AAB04638.1; JOINED.
 DR EMBL: M26272; AAB04638.1; JOINED.
 DR EMBL: M26273; AAB04638.1; JOINED.
 DR EMBL: M26274; AAB04638.1; JOINED.
 DR EMBL: M26275; AAB04638.1; JOINED.
 DR EMBL: M26276; AAB04638.1; JOINED.
 DR EMBL: M26277; AAB04638.1; JOINED.
 DR EMBL: M26278; AAB04638.1; JOINED.
 DR EMBL: M26279; AAB04638.1; JOINED.
 DR EMBL: M26280; AAB04638.1; JOINED.
 DR EMBL: M26281; AAB04638.1; JOINED.
 DR EMBL: M26282; AAB04638.1; JOINED.
 DR EMBL: M26283; AAB04638.1; JOINED.
 DR EMBL: M26284; AAB04638.1; JOINED.
 DR EMBL: M26285; AAB04638.1; JOINED.
 DR EMBL: M26286; AAB04638.1; JOINED.
 DR EMBL: M26287; AAB04638.1; JOINED.
 DR EMBL: M26288; AAB04638.1; JOINED.
 DR EMBL: M26289; AAB04638.1; JOINED.
 DR EMBL: M26290; AAB04638.1; JOINED.
 DR EMBL: M26291; AAB04638.1; JOINED.
 DR EMBL: M26292; AAB04638.1; JOINED.
 DR EMBL: M26293; AAB04638.1; JOINED.
 DR EMBL: M26294; AAB04638.1; JOINED.
 DR EMBL: M26295; AAB04638.1; JOINED.
 DR EMBL: M26296; AAB04638.1; JOINED.
 DR EMBL: M26297; AAB04638.1; JOINED.
 DR EMBL: M26298; AAB04638.1; JOINED.
 DR EMBL: M26299; AAB04638.1; JOINED.
 DR EMBL: M26300; AAB04638.1; JOINED.
 DR EMBL: M26301; AAB04638.1; JOINED.
 DR EMBL: M26302; AAB04638.1; JOINED.
 DR EMBL: M26303; AAB04638.1; JOINED.
 DR EMBL: M26304; AAB04638.1; JOINED.
 DR EMBL: M26305; AAB04638.1; JOINED.
 DR EMBL: M26306; AAB04638.1; JOINED.
 DR EMBL: M26307; AAB04638.1; JOINED.
 DR EMBL: M26308; AAB04638.1; JOINED.
 DR EMBL: M26309; AAB04638.1; JOINED.
 DR EMBL: M26310; AAB04638.1; JOINED.
 DR EMBL: M26311; AAB04638.1; JOINED.
 DR EMBL: M26312; AAB04638.1; JOINED.
 DR EMBL: M26313; AAB04638.1; JOINED.
 DR EMBL: M26314; AAB04638.1; JOINED.
 DR EMBL: M26315; AAB04638.1; JOINED.
 DR EMBL: M26316; AAB04638.1; JOINED.
 DR EMBL: M26317; AAB04638.1; JOINED.
 DR EMBL: M26318; AAB04638.1; JOINED.
 DR EMBL: M26319; AAB04638.1; JOINED.
 DR EMBL: M26320; AAB04638.1; JOINED.
 DR EMBL: M26321; AAB04638.1; JOINED.
 DR EMBL: M26322; AAB04638.1; JOINED.
 DR EMBL: M26323; AAB04638.1; JOINED.
 DR EMBL: M26324; AAB04638.1; JOINED.
 DR EMBL: M26325; AAB04638.1; JOINED.
 DR EMBL: M26326; AAB04638.1; JOINED.
 DR EMBL: M26327; AAB04638.1; JOINED.
 DR EMBL: M26328; AAB04638.1; JOINED.
 DR EMBL: M26329; AAB04638.1; JOINED.
 DR EMBL: M26330; AAB04638.1; JOINED.
 DR EMBL: M26331; AAB04638.1; JOINED.
 DR EMBL: M26332; AAB04638.1; JOINED.
 DR EMBL: M26333; AAB04638.1; JOINED.
 DR EMBL: M26334; AAB04638.1; JOINED.
 DR EMBL: M26335; AAB04638.1; JOINED.
 DR EMBL: M26336; AAB04

Best Local Similarity 29.4%; Pred. NO. 3e-06;
Matches 35; Conservative 16; Mismatches 62; Indels 6; Gaps 3;

Yy 7 PELGTMDCHPPGN---FSSSOCAFCSEGNTNLGIETTCGPFGMSPEPTCQVIG 63
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 853 PPKRTNGNHTGSNTIARSPGMSILYSODQGYLLVAGLALCTHGTWSGAPRCKEYNC 912

Qy 64 EPLSADL-GIMNCNPPLASFSTSNCFITCEGETELLIGKTKTICSESSINSPIQC 121
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 913 S--SPADMGGIGKCLEPRKMYYGAVVTLECDGYMLGSGPSOSCOSSHQNNPPLAVCR 969

RESULT 37

CFAA_HUMAN STANDARD: PRI: 764 AA.

ID CFAA_HUMAN STANDARD: PRI: 764 AA.
AC P00751; O15006; Q29944; Q9BTE5; Q9BX92;
DT 21-JUL-1986 (Rel. 01, Created)
DR 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Complement factor B precursor (EC 3.4.21.4) (C3/C5 convertase)
DE (Propeptid factor B) (glycine-rich beta glycoprotein) (GBG) (BBF2).
BF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
[1]
SEQUENCE FROM N.A. (ALLELES S. FA AND FB).
RX MEDLINE=91065702; PubMed=2249679;
RA Davinche C., Abbal M., Clerc A.;
RT "Molecular characterization of human complement factor B subtypes.";
RL Immunogenetics 32:309-312(1990).
RN [2]
SEQUENCE FROM N.A. (ALLELE S).
RC TISSUE=Liver:
RX MEDLINE=94237735; PubMed=8181962;
RA Mejia J.E., Jahn I., de la Salle H., Hauptmann G.;
RL "Human factor B. Complete cDNA sequence of the BF*S allele.";
RN Hum. Immunol. 39:49-53(1994).
RN [3]
SEQUENCE FROM N.A. (ALLELE S).
RC TISSUE=Liver:
MEDLINE=94041399; PubMed=8225386;
RX Schwaebel W., Luetlitz B., Sokolowski T., Estaller C., Weiss E.H.,
RA Meyer Zum Buschenfelde K.-H., Whaley K., Dipold W.;
RT "Human complement factor B: functional properties of a recombinant
RT zymogen of the alternative activation pathway convertase.";
RL Immunobiology 188:221-232(1993).
RN [4]
SEQUENCE FROM N.A. (ALLELE S).
RP MEDLINE=94067177; PubMed=8247029;
RX Horiochi T., Kim S., Matsumoto M., Watanabe I., Fujita S.,
RA Volanakis J.E.;
RT "Human complement factor B: cDNA cloning, nucleotide sequencing,
RT phenotypic conversion by site-directed mutagenesis and expression,";
RN Mol. Immunol. 30:1587-1592(1993).
RN [5]
SEQUENCE FROM N.A.
RP Rowen L., Dankers C., Baskin D., Faust J., Ioretz C., Ahern M.E.,
RA Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;
RT "Sequence determination of 300 kilobases of the human class III MHC
RT locus.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
SEQUENCE FROM N.A.
RP TISSUE=Colon;
RA Strausberg R.;
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
SEQUENCE FROM N.A. (ISOFORM 2).
RA Uasthen T., Kanerva J., Poutanen K.E., Saarinen-Pihkala U.,
RA Lokki M.-L.;
RT "Expression and alternative splicing of human factor B gene in

RT leukemic mononuclear cells.";
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
RN [8]
RP SEQUENCE OF 26-764, PARTIAL SEQUENCE FROM N.A., AND CARBOHYDRATES.
RX MEDLINE-84161997; PubMed-6546754;
RA Mole J.E., Anderson J.K., Davison E.A., Woods D.E.;
RT "Complete primary structure for the zymogen of human complement
RT factor B.";
RL J. Biol. Chem. 259:3407-3412(1984).
RN [9]
RP SEQUENCE OF 260-764.
RX MEDLINE-83204002; PubMed-6342610;
RA Christie D.L., Gagnon J.;
RT "Amino acid sequence of the Bb fragment from complement factor B.
RT Sequence of the major cyanogen bromide-cleavage peptide (CB-II) and
RL completion of the sequence of the Bb fragment.";
RL Biochem. J. 209:61-70(1983).
RN [10]
RP SEQUENCE OF 339-764 FROM N.A.
RX MEDLINE-83373641; PubMed-6308626;
RA Campbell R.D., Porter R.R.;
RT "Molecular cloning and characterization of the gene coding for human
RT complement protein factor B.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4464-4468(1983).
RN [11]
RP SEQUENCE OF 467-595 AND 752-764 FROM N.A.
RX MEDLINE-83039428; PubMed-6957884;
RA Woods D.E., Markham A.F., Ricker A.T., Goldberger G., Colten H.R.;
RT "Isolation of cDNA clones for the human complement protein factor B,
RT a class III major histocompatibility complex gene product.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:5661-5665(1982).
RN [12]
RP SEQUENCE OF 16-259 FROM N.A.
RX MEDLINE-84158524; PubMed-6323161;
RA Morley B.J., Campbell R.D.;
RT "Internal homologues of the Ba fragment from human complement
RT component factor B, a class III MHC antigen.";
RL EMBO J. 3:153-157(1984).
RN [13]
RP SEQUENCE OF 1-99 FROM N.A.
RX TISSUE-Blood;
RC MEDLINE-87102880; PubMed-3643061;
RA Wu L.C., Morley B.J., Campbell R.D.;
RT "Cell-specific expression of the human complement protein factor B
RT gene: evidence for the role of two distinct 5'-flanking elements.";
RL Cell 48:331-342(1987).
RN [14]
RP GLYCATION IN POSITION 291.
RX MEDLINE-91174758; PubMed-2006911;
RA Niemann M.A., Bhown A.S., Miller E.J.;
RT "The principal site of glycation of human complement factor B.";
RL Biochem. J. 274:473-480(1991).
CC -1- FUNCTION: FACTOR B WHICH IS PART OF THE ALTERNATE PATHWAY OF THE
CC COMPLEMENT SYSTEM IS CLEAVED BY FACTOR D INTO 2 FRAGMENTS: BA AND
CC BB. BA, A SERINE PROTEASE, THEN COMBINES WITH COMPLEMENT FACTOR 3B
CC TO GENERATE THE C3 OR C5 CONVERTASE. IT HAS ALSO BEEN IMPLICATED
CC IN PROLIFERATION AND DIFFERENTIATION OF PREACTIVATED B
CC LYMPHOCYTES, RAPID SPREADING OF PERIPHERAL BLOOD MONOCYTES,
CC STIMULATION OF LYMPHOCYTE BLASTOGENESIS AND LYSIS OF ERYTHROCYTES.
CC BA INHIBITS THE PROLIFERATION OF PREACTIVATED B LYMPHOCYTES.
CC -1- CATALYTIC ACTIVITY: Cleaves C3 in the alpha-chain to yield C3a and
CC C3b. Cleaves C5 in the alpha-chain to yield C5a and C5b. Both
CC cleavages take place at the C-terminal of an arginine residue.
CC -1- SUBUNIT: MONOMER.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- POLYMORPHISM: TWO MAJOR VARIANTS, F AND S, AND 2 MINOR VARIANTS,
CC AS WELL AS AT LEAST 14 VERY RARE VARIANTS, HAVE BEEN IDENTIFIED.
CC -1- MISCELLANEOUS: FACTOR B IS A MAJOR HISTOCOMPATIBILITY COMPLEX
CC CLASS-III PROTEIN.
CC -1- SIMILARITY: WITH COMPLEMENT C2.
CC -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 VWFA DOMAIN.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X72875; CAA51389.1; -
CC EMBL: S67310; AAD13989.1; -
CC EMBL: L15702; AAA16820.1; -
CC EMBL: X00284; CAA25077.1; -
CC EMBL: AF019413; AAB67977.1; -
CC EMBL: BC004143; AAH04143.1; -
CC EMBL: AF349679; AAK30167.1; -
CC EMBL: K01566; AAA36225.2; -
CC EMBL: J00125; -; NOT_ANNOTATED_CDS.
CC EMBL: J00126; AAA36226.1; -
CC EMBL: J00185; AAA36219.1; ALT_SEQ.
CC EMBL: J00186; AAA36220.1; -
CC EMBL: M15082; AAA59625.1; -
CC PIR: A00934; BBHU.
CC PIR: S14339; S14339.
CC PIR: S34075; S34075.
CC HSP: P00734; 2HNT.
CC MEROPS: S01.196; -.
CC SWISS-2DPAGE: P00751; HUMAN.
CC S1ena-2DPAGE: P00751; -.
CC MIM: 138470; -.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000436; Sush1_SCR_CCP.
CC InterPro: IPR001254; Trypsin.
CC InterPro: IPR002035; VWFA.
CC Pfam: PF00084; sush1; 3.
CC Pfam: PF00089; trypsin; 1.
CC Pfam: PF00092; vwa; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00453; VWFADOMAIN.
CC SMART: SM00032; CCP; 3.
CC SMART: SM00020; TRYP_SPE; 1.
CC SMART: SM00327; VWA; 1.
CC PROSITE: PS50240; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HTS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
CC PROSITE: PS50234; VWFA; 1.
CC Complement alternate pathway; Plasma; Hydrolase; Serine protease;
CC Glycoprotein; Repeat; Sush1; Signal; Polymorphism; Zymogen;
CC KW
CC Alternative splicing;
CC FT SIGNAL 1 25
CC FT CHAIN 26 764
CC FT CHAIN 26 259
CC FT CHAIN 260 764
CC FT CHAIN 36 99
CC FT DOMAIN 102 159
CC FT DOMAIN 164 219
CC FT DOMAIN 270 469
CC FT DOMAIN 482 764
CC FT ACT_SITE 526 536
CC FT ACT_SITE 576 576
CC FT ACT_SITE 639 699
CC FT ACT_SITE 37 76
CC FT DISULFID 103 145
CC FT DISULFID 131 158
CC FT DISULFID 165 205
CC
CC Query Match 19.2%; Score 140.5; DB 1; Length 764;
CC Best local Similarity 25.2%; Pred. No. 2.5e-06;
CC Matches 35; Conservative 14; Mismatches 55; Indels 35; Gaps 4;

```

QY 1 CEPLEAPELGTMDCTHPFGN-----FSSSCAFSCSGEGLNLGTGIETTCCGPGNW 51
DB 98 CRAIHCP-----PHDENGEYMRSPYNNVSDSEIFHCYDGYLRCSANTRCQVNGRW 151
OY 52 SSPEPTCOVIOCEPLASAPDLGIMNCSP-----LASFSTACPFICSEGLIGK 102
DB 152 SGOAIC-----DNGAGYCSNPGIPIGRKVGSOYRLDSEVYHCSRGJLTGRS 200
OY 103 KKTICSSGWSNPSPIQ 121
DB 201 ORRTQEGGWSWGTEPSCQ 219

RESULT 38
VCP_VACCV STANDARD; PRT; 263 AA.
ID VCP_VACCV
AC P10998;
DT 01-JUL-1989 (Rel. 11, Last Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Complement control protein precursor (VCP) (Secretory protein 35)
DE (Protein C3) (28 kDa protein).
OS C3L.
OS Vaccinia virus (strain WR), and
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10254, 10249;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 20-37.
RC STRAIN=WR;
RX MEDLINE=88318974; PubMed=3412473;
RA Kotwal G.J., Moss B.;
RT "Vaccinia virus encodes a secretory polypeptide structurally related
RL to complement control proteins.";
RL Nature 335:176-178(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WR;
RX MEDLINE=89073756; PubMed=2849238;
RA Kotwal G.J., Moss B.;
RT "Analysis of a large cluster of nonessential genes deleted from a
RL vaccinia virus terminal transposition mutant.";
RL Virology 167:524-537(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COPENHAGEN;
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
RT Paoletti E.;
RL "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [4]
RP COMPLETE GENOME.
RC STRAIN=COPENHAGEN;
RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
RT Paoletti E.;
RL "Appendix to 'The complete DNA sequence of vaccinia virus.'";
RL Virology 179:517-563(1990).
RN [5]
RP FUNCTION.
RX MEDLINE=92115714; PubMed=1731333;
RA Isaacs S.N., Kotwal G.J., Moss B.;
RT "antibody-virus complement-control protein prevents
RT antibody-dependent complement-enhanced neutralization of infectivity
RL and contributes to virulence.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:628-632(1992).
RN [6]
RP STRUCTURE BY NMR OF 146-263.
RX MEDLINE=97446168; PubMed=2929352;
RA Wiles A.P., Shaw G., Bright J., Perczel A., Campbell I.D.,
RT Barlow P.N.;
RT "NMR studies of a viral protein that mimics the regulators of

```

```

RT complement activation.";
RL J. Mol. Biol. 272:253-265(1997).
CC -I- FUNCTION: SERVES TO PROTECT THE VIRUS AGAINST COMPLEMENT ATTACK BY
CC INHIBITING BOTH CLASSICAL AND ALTERNATIVE PATHWAYS OF COMPLEMENT
CC ACTIVATION. BINDS C3B AND C4B.
CC -I- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
CC COMPLEMENT ACTIVATION (RCA).
CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13166; CAA31564.1; -;
DR EMBL; M22812; AAA69605.1; -;
DR EMBL; M35027; AAA47997.1; -;
DR PIR; A31005; WMVZSP.
DR PDB; 1VVC; 03-DEC-97.
DR PDB; 1VVD; 03-DEC-97.
DR PDB; 1VVE; 03-DEC-97.
DR InterPro; IPR000436; Sushi_SCR.CCP.
DR Pfam; PF00084; sushi; 4.
DR SMART; SM00032; CCP; 4.
FW Signal: Repeat: Sushi; 3D-structure.
FT CHAIN 1
FT SIGNAL 19
FT DOMAIN 20 263 COMPLEMENT CONTROL PROTEIN.
FT DOMAIN 20 82 SUSHI 1.
FT DOMAIN 85 144 SUSHI 2.
FT DOMAIN 147 202 SUSHI 3.
FT DOMAIN 205 262 SUSHI 4.
FT DISULFID 21 70 BY SIMILARITY.
FT DISULFID 54 81 BY SIMILARITY.
FT DISULFID 86 126 BY SIMILARITY.
FT DISULFID 112 143 BY SIMILARITY.
FT DISULFID 148 190 BY SIMILARITY.
FT DISULFID 176 201 BY SIMILARITY.
FT DISULFID 206 248 BY SIMILARITY.
FT DISULFID 234 261 BY SIMILARITY.
SQ SEQUENCE 263 AA; 28629 MW; E4322CC9A6BEF8997 CRC64;

Query Match 19.0%; Score 139; DB 1; Length 263;
Best Local Similarity 33.1%; Pred. No. 1.1e-06;
Matches 39; Conservative 17; Mismatches 46; Indels 16; Gaps 8;

OY 10 GTMDCTHPFGNFFSSSCAFSCSGEGLNLGTGIETTCC--GPFGN--SSPEPTCOVIOCEP 65
DB 95 GOLD-----IGVDPRGSSITFYSCNSGYHLIGESMSYCELGSTGSMVWMPAPICSEYKQ 149

OY 66 LSAPDLGIMNCSPHLASFST--SACFFICSEGLIGKKTICSESSGWSNPSPIQ 121
DB 150 -SP- -SISNGRHNGYEDFYTDGSSVYVYSCNSGYLIGNSGVLC -SGGEWSDP-PTCQ 202

RESULT 39
ID DAF_HUMAN STANDARD; PRT; 381 AA.
AC P08174; P09679; P78361;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Complement decay-accelerating factor precursor (CD55 antigen).
GN DAF OR CR OR CD55.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

```

RA MEDLINE=87115845; PubMed=2433596;
RA Caras I.W., Davitz M.A., Rhee L., Meddell G., Martin D.W. Jr.,
RA Nussenzweig V.;
RT "Cloning of decay-accelerating factor suggests novel use of splicing
RT to generate two proteins.";
RL Nature 325:545-549(1987).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 6-381 FROM N.A. (ISOFORM 2).
RX MEDLINE=871175602; PubMed=2436222;
RA Medof M.E., Lublin D.M., Holers V.M., Ayers D.J., Getty R.R.,
RA Leykam J.F., Atkinson J.P., Tykocinski M.L.;
RT "Cloning and characterization of cDNAs encoding the complete sequence
RT of decay-accelerating factor of human complement.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2007-2011(1987).
RN [4]
RN SEQUENCE OF 35-381 FROM N.A. (ISOFORM 2).
RC TISSUE=Hippocampus;
RA Kumar V.B., Hyung C., Nakra R., Walters M., Sasser T., Bernardo A.;
RT "Decay-accelerating factor (DAF; CD 55) in the brain of Alzheimer's
RT disease patients.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE OF 1-100 FROM N.A.
RX MEDLINE=91271256; PubMed=1711208;
RA Ewulonu U.K., Ravi L., Medof M.E.;
RT "Characterization of the decay-accelerating factor gene promoter
RT region.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:4675-4679(1991).
RN [6]
RN SEQUENCE OF 35-46.
RC TISSUE=Urine;
RX MEDLINE=91291869; PubMed=1712233;
RA Nakano Y., Sugita Y., Ishikawa Y., Choi N.-H., Tobe T., Tomita M.;
RT "Isolation of two forms of decay-accelerating factor (DAF) from human
RT urine.";
RL Biochim. Biophys. Acta 1074:326-330(1991).
RN [7]
RN GPI-ANCHOR.
RX MEDLINE=91093238; PubMed=1824699;
RA Moran P., Raab H., Kohr W.J., Caras I.W.;
RT "Glycophospholipid membrane anchor attachment. Molecular analysis of
RT the cleavage/attachment site.";
RL J. Biol. Chem. 266:1250-1257(1991).
RN [8]
RN DISULFIDE BONDS IN SUSHI DOMAINS.
RX MEDLINE=92305034; PubMed=1377029;
RA Nakano Y., Sumida K., Kikuta N., Miura N.-H., Tobe T., Tomita M.;
RT "Complete determination of disulfide bonds localized within the short
RT consensus repeat units of decay accelerating factor (CD55 antigen).";
RL Biochim. Biophys. Acta 1116:235-240(1992).
RN [9]
RN FUNCTION AS A ECHOVIRUS RECEPTOR.
RX MEDLINE=95045399; PubMed=7525274;
RA Ward T., Pipkin P.A., Clarkson N.A., Stone D.M., Minor P.D.,
RA Almond J.W.;
RT "Decay-accelerating factor CD55 is identified as the receptor for
RT echovirus 7 using CELICs, a rapid immuno-focal cloning method.";
RL EMBO J. 13:5070-5074(1994).
RN [10]
RN VARIANT BLOOD GROUP DR(A-).
RX MEDLINE=94325573; PubMed=7519480;
RA Lublin D.M., Mallinson G., Poole E., Reid M.E., Thompson E.S.,
RA Ferdman B.R., Telen M.J., Anstee D.J., Tanner M.J.A.;
RT "Molecular basis of reduced or absent expression of
RT decay-accelerating factor in Cromer blood group phenotypes.";
RL Blood 84:1276-1282(1994).
CC -1- FUNCTION: THIS PROTEIN RECOGNIZES C4B AND C3B FRAGMENTS THAT
CC CONDENSE WITH CELL-SURFACE HYDROXYL OR AMINO GROUPS WHEN NASCENT

CC C4B AND C3B ARE LOCALLY GENERATED DURING C4 AND C3 ACTIVATION.
CC INTERACTION OF DAF WITH CELL-ASSOCIATED C4B AND C3B POLYPEPTIDES
CC INTERFERES WITH THEIR ABILITY TO CATALYZE THE CONVERSION OF C2 AND
CC FACTOR B TO ENZYMATICALLY ACTIVE C2A AND Bb AND THEREBY PREVENTS
CC THE FORMATION OF C4B2A AND C3BB. THE AMPLIFICATION CONVERSASES OF
CC THE COMPLEMENT CASCADE.
CC -1- FUNCTION: ALSO ACTS AS THE RECEPTOR FOR ECHOVIRUS 7 AND RELATED
CC VIRUSES (ECHOVIRUSES 13, 21, 29 AND 33).
CC -1- SUBUNIT: MONOMER (MAJOR FORM) AND NON-DISULFIDE-LINKED, COVALENT
CC HOMODIMER (MINOR FORM).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, 1/DAF-1 AND 2/DAF-2 (SHOWN
CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED ON THE PLASMA MEMBRANES OF ALL CELL
CC TYPES THAT ARE IN INTIMATE CONTACT WITH PLASMA COMPLEMENT
CC PROTEINS. IT IS ALSO FOUND ON THE SURFACES OF EPITHELIAL CELLS
CC Lining EXTRACELLULAR COMPARTMENTS, AND VARIANTS OF THE MOLECULE
CC ARE PRESENT IN BODY FLUIDS AND IN EXTRACELLULAR MATRIX.
CC -1- DOMAIN: THE FIRST SUSHI DOMAIN (SCR1) IS NOT NECESSARY FOR
CC FUNCTION. SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE
CC ACTIVE SITE ON SCR3 (BY SIMILARITY).
CC -1- PTR: THE SER/THR-RICH DOMAIN IS HEAVILY O-GLYCOSYLATED.
CC -1- POLYMORPHISM: DAF IS RESPONSIBLE FOR THE CROMER BLOOD GROUP
CC SYSTEM. IT CONSISTS OF AT LEAST SEVEN HIGH-INCIDENCE (CR(A),
CC TC(A), DR(A), ES(A), WES(B), UMC, AND IFC) AND LOW-INCIDENCE
CC (TC(B), TC(C), AND WES(A)) ANTIGENS THAT RESIDE ON DAF. IN THE
CC CROMER PHENOTYPES DR(A-) AND INAB THERE IS REDUCED OR ABSENT
CC EXPRESSION OF DAF, RESPECTIVELY. IN THE CASE OF THE DR(A-) PHENOTYPE,
CC A SINGLE NUCLEOTIDE SUBSTITUTION WITHIN EXON 5 ACCOUNTS
CC FOR TWO CHANGES: A SIMPLE AMINO ACID SUBSTITUTION THAT IS THE
CC BASIS OF THE ANTIGENIC VARIATION, AND AN ALTERNATIVE SPLICING
CC EVENT THAT UNDERLIES THE DECREASED EXPRESSION OF DAF IN THIS
CC PHENOTYPE.
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
CC (RCA) FAMILY.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD55 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd55.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M31516; AAA52169.1; -;
CC DR EMBL; M30142; AAA52168.1; -;
CC DR EMBL; BC001288; AAH01288.1; -;
CC DR EMBL; M15799; AA52167.1; -;
CC DR EMBL; U88576; AAB48622.1; -;
CC DR EMBL; M64653; AAA52170.1; -;
CC DR EMBL; M64356; AAA52170.1; JOINED.
CC DR EMBL; S72858; AAC60633.1; -;
CC DR PIR; B26359; B26359.
CC DR PIR; A26359; A26359.
CC DR PIR; S16187; S16187.
CC DR PIR; A39101; A39101.
CC DR PIR; S23138; S23138.
CC DR HSSP; P08603; 1HCC.
CC DR MIM; 125240; -;
CC DR InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00084; sushi_4.
CC SMART; SM00032; CCP; 4.
CC KW Complement pathway; Plasma; glycoprotein; Membrane; Repeat;
CC KW Alternative splicing; GPI-anchor; Signal; Sushi; Polymorphism;
CC Blood group antigen.
CC SIGNAL 1 34
CC CHAIN 35 353 COMPLEMENT DECAY-ACCELERATING FACTOR.
CC PROPEP 35 381 REMOVED IN MATURE FORM.
CC DOMAIN 35 95 SUSHI 1.
CC FT 97 159 SUSHI 2.

```

FT DOMAIN 162 221 SUSHI 3.
FT DOMAIN 224 284 SUSHI 4.
FT DOMAIN 287 356 SER/THR-RICH.
FT DISULFID 65 81
FT DISULFID 98 145
FT DISULFID 129 158
FT DISULFID 163 204
FT DISULFID 190 220
FT DISULFID 225 267
FT DISULFID 253 283
FT CARBOHYD 95 95
FT LIPID 353 353
FT VARSPPLIC 362 381

N-LINKED (GLCNAC. . .) (POTENTIAL).
GPI-ANCHOR.
HTCFTLLGLGLTVMGLLT -> SRPVTOAGNRKCDRSSL
OSRTPGKRSFHFSLPSSWYFAHVEHDFRFDASDHGLA
DLAKEELRKYTOVLEFLVS (IN ISOFORM 1).
R -> L (IN TC(B) ANTIGEN).
/FTID-VAR_001997.
R -> P (IN TC(C) ANTIGEN).
/FTID-VAR_001998.
L -> R (IN MES(A) ANTIGEN).
/FTID-VAR_001999.
S -> L (IN DR(A-) ANTIGEN).
/FTID-VAR_002000.
A -> P (IN CR(A-) ANTIGEN).
/FTID-VAR_002001.
T -> I (IN REF. 1, 2 AND 4).
S -> M (IN REF. 3).
S -> T (IN REF. 4).
Q -> H (IN REF. 4).
FT CONFLICT 187 187
FT CONFLICT 297 297
SQ SEQUENCE 381 AA; 41388 MW; 29138EB6B4B565E CRC64;

```

Query Match 19.0%; Score 139; DB 1; Length 381;
 Best Local Similarity 29.4%; Pred. No. 1.7e-06;
 Matches 32; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

```

QY 19 GNFSSQAFSCSEGTNLGIEETTC--GPRGMSPPPTCOVIOCEPLSAPDLGIMN 75
    | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 178 GILFGATISFCNTGKLPSTSSFLIGSSVQWMDPLPECEIYCPAPQIDNIIQ 237
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 76 CSPLASFTSACTFICSEGTLEIKKKKTTIC--ESSGIWSNPSPIQ 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 GERD--HYGYROSIVTYACNKGFTMIGHSITCYVNDGEMSGPPPCR 284
    : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 40
 DAF_PONPY STANDARD; PRT; 340 AA.

```

AC P49457;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Complement decay-accelerating factor (CD55) (Fragment).
DAF OR CD55.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94110622; PubMed=7506731;
RA Nickells M.W., Alvarez J.I., Lublin D.M., Atkinson J.P.;
RT "Characterization of DAF-2, a high molecular weight form of decay-
RT accelerating factor (DAF, CD55), as a covalently cross-linked dimer
RT of DAF-1."
RL J. Immunol. 152:676-685(1994).
CC -1- FUNCTION: THIS PROTEIN RECOGNIZES C4B AND C3B FRAGMENTS THAT
CC CONDENSE WITH CELL-SURFACE HYDROXYL OR AMINO GROUPS WHEN NASCENT
CC C4B AND C3B ARE LOCALLY GENERATED DURING C4 AND C3 ACTIVATION.
CC INTERACTION OF DAF WITH CELL-ASSOCIATED C4B AND C3B POLYPEPTIDES
CC INTERFERES WITH THEIR ABILITY TO CATALYZE THE CONVERSION OF C2 AND
CC FACTOR B TO ENZYMATICALLY ACTIVE C2A AND BB AND THEREBY PREVENTS

```

```

CC THE FORMATION OF C4B2A AND C3BB, THE AMPLIFICATION CONVERTASES OF
CC THE COMPLEMENT CASCADE (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (MAJOR FORM) AND NON-DISULFIDE-LINKED, COVALENT
CC HOMODIMER (MINOR FORM).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF DAF (DAF-2, SHOWN HERE, AND
CC DAF-1) ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
CC -1- FUNCTION: SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE
CC ACTIVE SITE ON SCR3 (BY SIMILARITY).
CC -1- PTM: THE SER/THR-RICH DOMAIN IS HEAVILY O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
CC (RCA) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S67775; AAC60609.1; -.
DR HSSP; P08603; IHFI.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 4.
DR SMART; SM00032; CCP; 4.
KW Complement pathway; Plasma; Glycoprotein; Membrane; Repeat;
KW Alternative splicing; GPI-anchor; Sush1.
FT NON-TER 1
FT CHAIN <1 312 COMPLEMENT DECAY-ACCELERATING FACTOR.
FT PROPEP 313 340 REMOVED IN NATURE FORM (BY SIMILARITY).
FT DOMAIN <1 54 SUSHI 1.
FT DOMAIN 56 118 SUSHI 2.
FT DOMAIN 121 180 SUSHI 3.
FT DOMAIN 183 243 SUSHI 4.
FT DOMAIN 246 315 SER/THR-RICH.
FT DISULFID 24 53 BY SIMILARITY.
FT DISULFID 57 104 BY SIMILARITY.
FT DISULFID 88 117 BY SIMILARITY.
FT DISULFID 122 163 BY SIMILARITY.
FT DISULFID 149 179 BY SIMILARITY.
FT DISULFID 184 226 BY SIMILARITY.
FT DISULFID 212 242 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 312 312 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 340 AA; 37180 MW; D3D865C058204290 CRC64;

```

Query Match 18.7%; Score 137; DB 1; Length 340;
 Best Local Similarity 29.5%; Pred. No. 2.3e-06;
 Matches 31; Conservative 16; Mismatches 50; Indels 8; Gaps 3;

```

QY 23 FSSQAFSCSEGTNLGIEETTC--GPRGMSPPPTCOVIOCEPLSAPDLGIMNCSHP 79
    | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 141 FGATISFCNTGKLPSTSSFLIGSSVQWMDPLPECEIYCPAPQIDNIIQGRND 200
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 80 LASFTSACTFICSEGTLEIKKKKTTIC--ESSGIWSNPSPIQ 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 --HYGYROSIVTYACNKGFTMIGHSITCYVNDGEMSGPPPCR 243
    : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: September 7, 2002, 10:23:46
 Job time: 539 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:15:28 : Search time 36.49 Seconds
(without alignments)
347.596 Million cell updates/sec

Title: US-09-119-209-2_COPY_197_328

Perfect score: 733
Sequence: 1 CEPLEAPDELGTMDCTHPEFGN.....WSNPSICQKDKDSFSMIKE 132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	99.2	385	1 A34015	L-selectin precurs
2	690	94.1	372	2 JC5377	L-selectin precurs
3	599	81.7	376	2 JC4892	L-selectin precurs
4	537	73.3	370	2 S22124	L-selectin precurs
5	535	73.0	372	2 S23936	L-selectin precurs
6	518	70.7	323	1 S09702	L-selectin precurs
7	518	70.7	372	1 A32375	L-selectin precurs
8	315	43.0	830	2 A30359	P-selectin precurs
9	313	42.7	768	2 I53821	P-selectin - rat
10	305.5	41.7	612	2 B42755	E-selectin precurs
11	300	40.9	768	2 A42755	P-selectin precurs
12	296	40.4	610	2 A35045	E-selectin precurs
13	290	39.6	646	2 JN0473	P-selectin precurs
14	288	39.3	551	2 I46709	P-selectin precurs
15	271.5	37.0	485	2 S36772	endothelial leukoc
16	245	33.4	449	2 JC5092	E-selectin - bovin
17	173	23.6	442	1 NBH0HS	E-selectin - pig
18	173	23.6	1231	1 NBH0H	complement factor
19	167.5	22.9	2014	2 I36936	complement factor
20	163.5	22.3	2489	2 I73012	complement recepto
21	160	21.8	597	1 S53711	complement C3b/C4b
22	155.5	21.2	1019	2 A38738	C4BP alpha chain p
23	155.5	20.5	560	2 T16833	complement factor
24	150.5	20.5	1025	1 A43526	coagulation factor
25	150.5	20.5	1025	1 A43526	hypothetical prote
26	149.5	20.4	597	1 NBH0C4	complement C3d/Eps
27	148	20.2	363	2 B45900	C4b-binding protel
28	148	20.2	676	2 A45900	complement C3b rec
29	147.5	20.1	482	2 A34924	complement C3b/C4b

30	147.5	20.1	669	2 S65551	factor H - bovine
31	147.5	20.1	868	2 T20239	hypothetical prote
32	147.5	20.1	1927	2 T34288	hypothetical prote
33	146	19.9	558	2 S57953	C4BP protein alpha
34	145	19.8	610	1 I46001	C4b-binding protel
35	144.5	19.7	469	1 NBMS04	C4BP protein beta
36	144	19.6	258	2 S57960	apolipoprotein H p
37	141.5	19.3	345	1 NBBO	complement C3d/Eps
38	141	19.2	1091	1 PL0009	complement factor
39	140.5	19.2	764	1 BBH0	apolipoprotein H h
40	139	19.0	263	1 MWV2SP	decay-accelerating
41	139	19.0	381	1 B26359	decay-accelerating
42	139	19.0	440	2 A26359	complement control
43	138.5	18.9	263	1 C36838	B18b protein - var
44	138.5	18.9	263	2 B72152	hypothetical prote
45	138.5	18.9	263	2 T28450	

ALIGNMENTS

RESULT 1
A34015
L-selectin precursor, long splice form - human
N:Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukoc
ral lymph node homing receptor Leu-8
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: I55333; S06798; J0104; A34015; A33912
R:Ord, D.C.; Ernst, T.J.; Zhou, L.J.; Rambaldi, A.; Sperlant, O.; Griffin, J.; Tedder
J. Biol. Chem. 265, 7760-7767, 1990
A:Title: Structure of the gene encoding the human leukocyte adhesion molecule-1 (LQ1,
A:Reference number: I55333; M01D:90243637
A:Accession: I55333
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 14-385 <ORF>
A:Cross-references: GB:X16214; NID:9187259; PIDN:AAB60700.1; PID:9386860
R:Cammerl, D.; James, S.P.; Stamenkovic, I.; Seed, B.
Nature 342, 78-82, 1989
A:Title: Leu-8/7Q1 is the human equivalent of the Mel-14 lymph node homing receptor.
A:Reference number: S06798; M01D:90044046
A:Accession: S06798
A:Molecule type: mRNA
A:Residues: 1-225, 'S', 227-385 <CAM>
A:Cross-references: EMBL:X17519; NID:934344; PIDN:CAB43536.1; PID:94902829
A:Note: This translation is not annotated in GenBank entry HSLF08, release 111.0
R:Tedder, T.F.; Isaacs, C.M.; Ernst, T.J.; Demetri, G.D.; Adler, D.A.; Distche, C.M.
J. Exp. Med. 170, 123-133, 1989
A:Title: Isolation and chromosomal localization of cDNAs encoding a novel human lymph
Oleins.
A:Reference number: J0104; M01D:89310350
A:Accession: J0104
A:Molecule type: mRNA
A:Residues: 1-230, 'N', 232, 'N', 234-254, 'E', 256-385 <TRD>
A:Cross-references: GB:X16150; NID:934428; PIDN:CAA34275.1; PID:934429
A:Note: The translated sequence in GenBank entry HSLVAM1, release 111.0, differs from
J. Bowen, B.R.; Nguyen, T.; Lasky, L.A.
J. Cell Biol. 109, 421-427, 1989
A:Title: Characterization of a human homologue of the murine peripheral lymph node hc
A:Reference number: A34015; M01D:89308881
A:Accession: A34015
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 14-49, 'Y', 51-190, 'H', 192-205, 'L', 207-226, 'F', 228-385 <BOW>
A:Cross-references: GB:X16070; NID:938092; PIDN:CAA34203.1; PID:938093
R:Siegelman, M.H.; Weissman, I.L.
Proc Natl. Acad. Sci. U.S.A. 86, 5562-5566, 1989
A:Title: Human homologue of mouse lymph node homing receptor: evolutionary conservati
A:Reference number: A33912; M01D:89315837
A:Accession: A33912
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 14-205, 'L', 207-385 <SIE>
A:Cross-references: GB:M25280; NID:9187182; PIDN:AAC63053.1; PID:9307134
C:Comment: For an alternative splice form, see PIR:S09702.
C:Genetics:
A:Gene: GDB:SELL; GDB:LINR; LSEL; LAM1; LYAM1; LAM-1
A:Cross-references: GDB:120157; GDB:118834; OMIM:153240
A:Map position: 1q22-1q23
A:Introns: 14/3; 42/1; 171/1; 207/1; 269/1; 331/1; 374/1; 380/2
C:Function:
A:Description: binds with low affinity to oligosaccharides like heparan sulfate and sialyl ment of leucocytes to areas of inflammation, and with CD16 mediates neutrophil-neutroph C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology; C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammation
F:1-51/Domain: signal sequence #status predicted <SIG>
F:42-168/Domain: C-type lectin homology <LCH>
F:52-385/Product: L-selectin #status predicted <MAT>
F:52-343/Domain: extracellular #status predicted <EXT>
F:173-204/Domain: EGF homology <EGF>
F:210-267/Domain: complement factor H repeat homology <FH1>
F:272-329/Domain: complement factor H repeat homology <FH2>
F:344-368/Domain: transmembrane #status predicted <TM>
F:369-385/Domain: intracellular #status predicted <INT>
F:73,117,190,245,259/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:377,380/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 99.2%; Score 727; DB 1; Length 385;
Best Local Similarity 99.2%; Pred. No. 1,2e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEPLAPLGLTMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 210 CEPLAPLGLTMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 269
Qy 61 IOCEPLAPDLGIMNCSPHPLASFSTACTFCISGTELGKTKTICSSGIMNSPSPIC 120
Db 270 IOCEPLAPDLGIMNCSPHPLASFSTACTFCISGTELGKTKTICSSGIMNSPSPIC 329
Qy 121 OKLDRSFSMIKE 132
Db 330 OKLDRSFSMIKE 341

RESULT 2
JC5377
L-selectin precursor - hamadryas baboon
C:Species: Papio hamadryas (hamadryas baboon)
C>Date: 02-Jun-1997 #sequence_revision 12-Sep-1997 #text_change 21-Jan-2000
C:Accession: JC5377; PC4315
R:Tsushita, N.; Fu, H.; Berg, E.L.
Gene 181, 219-220, 1996
A:Title: PCR cloning of the cDNA encoding baboon L-selectin.
A:Reference number: JC5377; MUID:97128794
A:Accession: JC5377
A:Molecule type: mRNA
A:Residues: 1-372 <TSU1>
A:Cross-references: GB:U52074; NID:91326148; PIDN:AA640903.1; PID:91326149
A:Accession: PC4315
A:Molecule type: protein
A:Residues: 37-43;142-148 <TSU2>
C:Comment: This receptor is involved in the initial adhesive interaction between lymphoc sites of inflammation.
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;
F:1-38/Domain: signal sequence #status predicted <SIG>
F:29-155/Domain: C-type lectin homology <LCH>
F:39-372/Product: L-selectin #status predicted <MAT>
F:39-157/Domain: calcium-binding #status predicted <CB>
F:160-191/Domain: EGF homology <EGF>
F:197-254/Domain: complement factor H repeat homology <FH1>
F:259-316/Domain: complement factor H repeat homology <FH2>
F:333-355/Domain: transmembrane #status predicted <TM>
F:356-372/Domain: intracellular #status predicted <INT>

Query Match 94.1%; Score 690; DB 2; Length 372;
Best Local Similarity 92.4%; Pred. No. 2.3e-54;
Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CEPLAPLGLTMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 197 CEPLAPLGLTMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 256
Qy 61 IOCEPLAPDLGIMNCSPHPLASFSTACTFCISGTELGKTKTICSSGIMNSPSPIC 120
Db 257 IOCEPLAPDLGIMNCSPHPLASFSTACTFCISGTELGKTKTICSSGIMNSPSPIC 316
Qy 121 OKLDRSFSMIKE 132
Db 317 OKLDRSFSMIKE 328

RESULT 3
JC4892
L-selectin precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 21-Jan-2000
C:Accession: JC4892
R:Qian, J.; Huang, X.; Marks, R.M.
Biochem. Biophys. Res. Commun. 225, 406-412, 1996
A:Title: Cloning of the cDNA for rabbit L-selectin and expression of recombinant prot A:Reference number: JC4892; MUID:96354800
A:Accession: JC4892
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-376 <OIA>
A:Cross-references: GB:U26535; NID:9847787; PIDN:AA67896.1; PID:9847788
C:Comment: This protein involved in leukocyte-endothelial adhesion; it mediates adhe F:1-37/Domain: signal sequence #status predicted <SIG>
F:29-155/Domain: C-type lectin homology; complement factor H repeat homology
F:38-376/Product: L-selectin #status predicted <MAT>
F:160-191/Domain: EGF homology <EGF>
F:197-254/Domain: complement factor H repeat homology <FH1>
F:259-316/Domain: complement factor H repeat homology <FH2>

Query Match 81.7%; Score 599; DB 2; Length 376;
Best Local Similarity 80.3%; Pred. No. 3.1e-46;
Matches 106; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

Qy 1 CEPLAPLGLTMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 197 CEPLAPLGLTMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 256
Qy 61 IOCEPLAPDLGIMNCSPHPLASFSTACTFCISGTELGKTKTICSSGIMNSPSPIC 120
Db 257 IOCEPLAPDLGIMNCSPHPLASFSTACTFCISGTELGKTKTICSSGIMNSPSPIC 316
Qy 121 OKLDRSFSMIKE 132
Db 317 OKLDRSFSMIKE 328

RESULT 4
S22124
L-selectin precursor - bovine
N:Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)
C:Species: Bos primigenius taurus (cattle)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jan-2000
C:Accession: S22124; A46531
R:Bosworth, B.T.
submitted to the EMBL Data Library, October 1991
A:Reference number: S22123
A:Accession: S22124
A:Status: preliminary
A:Molecule type: mRNA

Oy 1 CEPLAPELLGTMDCTHHPGNGFSSQCAFSCSEGTNLGIEETTCGPGFGNMSSPEPTCOV 60
|:||||| ||||| ||||| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 262 CQSLFAPLHGHTMDCTHPLAFAVYDSSCKRECPGYRMGRSDILHCHDSQWSEPLPTCA 321
Oy 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTLEIGKKTKICSSGIGMNSPSPIC 120
|:||||| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 322 IACEPLSEPLHSGMDCFPSTGAFGYNSCTCFRCEGTGFLMGNDAIHCADLQGMATAPAPVC 381
Oy 121 OKL 123
|:|
Db 382 EAL 384

RESULT 10
B42755
E-selectin precursor - mouse
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-1993 #sequence_revision 31-Dec-2000 #text_change 31-Dec-2000
C:Accession: S23174; B42755
R:Becker-Andre, M.; van Huijsduijnen, R.H.; Losberger, C.; Whelan, J.; Delamarier, J.F.
Eur. J. Biochem. 206, 401-411, 1992
A:Title: Murine endothelial leukocyte adhesion molecule 1 is a close structural and func
A:Reference number: S23174; MUID:92283265
A:Accession: S23174
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <BCE>
A:Cross-references: GB:M60778; NID:9193014; PIDN:AAA37547.1; PID:9193015
R:Weller, A.; Isenmann, S.; Vestweber, D.
J. Biol. Chem. 267, 15176-15183, 1992
A:Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-select
A:Reference number: A42755; MUID:92340571
A:Accession: B42755
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'MKATAGV', 1-389, 391-612 <MBL>
A:Cross-references: GB:M87862; NID:9193107
A:Experimental source: endothelial cells
A:Note: sequence extracted from NCBI backbone (NCBI:109470)
A:Note: It is uncertain in Genbank entry M87862, release 117.0, (PIDN:AAA37577.1; PID:91
A:Note: It is uncertain whether the Initiator is Met-1 or the AUG codon preceding that
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement fact
C:Keywords: glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:12-138/Domain: signal sequence #status predicted <SIG>
F:12-138/Domain: C-type lectin homology <LCH>
F:12-612/Product: P-selectin #status Predicted <MAT>
F:143-174/Domain: EGF homology <EGF>
F:180-238/Domain: complement factor H repeat homology <FH1>
F:243-300/Domain: complement factor H repeat homology <FH2>
F:305-363/Domain: complement factor H repeat homology <FH3>
F:366-426/Domain: complement factor H repeat homology <FH4>
F:431-489/Domain: complement factor H repeat homology <FH5>
F:494-548/Domain: complement factor H repeat homology <FH6>
F:25,391,528/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.7%; Score 305.5; DB 2; Length 612;
Best Local Similarity 42.6%; Pred. No. 8.2e-20;
Matches 52; Conservative 22; Mismatches 47; Indels 1; Gaps 1;

Oy 1 CEPLAPELLGTMDCTHHPGNGFSSQCAFSCSEGTNLGIEETT-CGPGFGNMSSPEPTCO 59
|:||||| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 180 CKPQHPDGLGIMNCSPHPLASFSTACTFICSEGTLEIGKKTKICSSGIGMNSPSPIC 239
Oy 60 VIOCEPLSAPDLGIMNCSPHPLASFSTACTFICSEGTLEIGKKTKICSSGIGMNSPSPIC 119
|:||||| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 240 VECALTFHPAIGIRKSSNPGSYPMWNTCTCFDCEGVRARVCAQNLQCTSSSIIWNEPFS 299
Oy 120 CQ 121
|:|
Db 300 CK 301

RESULT 11

P-selectin precursor - mouse

N:Alternate names: CD62; granule membrane protein 140; PADGEM

C:Species: Mus musculus (house mouse)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 19-May-2000

C:Accession: A42755; A4899

R:Weller, A.; Isenmann, S.; Vestweber, D.

J. Biol. Chem. 267, 15176-15183, 1992

A:Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-sele

A:Reference number: A42755; MUID:92340571

A:Accession: A42755

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-768 <MBL>

A:Cross-references: GB:M87861; NID:9200552; PIDN:AAA40008.1; PID:9200553

A:Experimental source: endothelial cells

A:Note: sequence extracted from NCBI backbone (NCBI:109467)

R:Sanders, W.E.; Wilson, R.W.; Ballantyne, C.M.; Beaudet, A.L.

Blood 80, 795-800, 1992

A:Title: Molecular cloning and analysis of in vivo expression of murine P-selectin.

A:Reference number: A4899; MUID:92345617

A:Accession: A4899

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-723, 'E', 725-768 <SAN>

A:Cross-references: GB:W7232; NID:9193565; PIDN:AAA37712.1; PID:9193566

A:Note: sequence extracted from NCBI backbone (NCBI:109900)

C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;

C:Keywords: cell adhesion; glycoprotein; phospholipid; phosphoprotein; transmembr

F:1-41/Domain: signal sequence #status predicted <SIG>

F:42-768/Product: P-selectin #status predicted <MAT>

F:163-194/Domain: EGF homology <EGF>

F:200-257/Domain: complement factor H repeat homology <FH01>

F:262-319/Domain: complement factor H repeat homology <FH02>

F:324-381/Domain: complement factor H repeat homology <FH03>

F:386-443/Domain: complement factor H repeat homology <FH04>

F:448-505/Domain: complement factor H repeat homology <FH05>

F:510-567/Domain: complement factor H repeat homology <FH06>

F:580-637/Domain: complement factor H repeat homology <FH07>

F:643-699/Domain: complement factor H repeat homology <FH08>

F:710-733/Domain: transmembrane #status predicted <TMN>

F:734-768/Domain: intracellular #status predicted <INT>

F:45,54,107,212,347,398,456,467,603,654,661,679/Binding site: carbohydrate (Asn) (cov

Query Match 40.9%; Score 300; DB 2; Length 768;
Best Local Similarity 42.3%; Pred. No. 3.2e-19;
Matches 52; Conservative 17; Mismatches 54; Indels 0; Gaps 0;

Oy 1 CEPLAPELLGTMDCTHHPGNGFSSQCAFSCSEGTNLGIEETTCGPGFGNMSSPEPTCOV 60
|:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 200 CKKVINPQHVHVNCCSPHPLASFSTACTFICSEGTLEIGKKTKICSSGIGMNSPSPIC 259
Oy 61 IOCEPLSAPDLGIMNCSPHPLASFSTACTFICSEGTLEIGKKTKICSSGIGMNSPSPIC 120
|:||||| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 260 VDCQSLFAPLHGHTMACMHPITAFYDSSCKRECPGYRARGSNITLHCTGSGGMSPEPLPTC 319

Oy 121 OKL 123
|:|
Db 320 EAL 322

RESULT 12

E-selectin precursor - human

N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1); LECAM-2

C:Species: Homo sapiens (man)

C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 21-Jul-2000

C:Accession: A38615; A35046; A32606

R:Collins, T.; Williams, A.; Johnston, G.I.; Kim, J.; Eddy, R.; Shows, T.; Gimbrone J

J. Biol. Chem. 266, 2466-2473, 1991

A>Title: Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule-1
A:Reference number: A38615; MUID:91115870

A:Accession: A38615

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-610 <COL>

A:Cross-references: GB:M61893; GB:M58017; NID:g182043; PIDN:AA52375.1; PID:g182046

R:Hession, C.; Osborn, L.; Goff, D.; Chi-Rosso, G.; Vassallo, C.; Psek, M.; Pittack, C.
Proc. Natl. Acad. Sci. U.S.A. 87, 1673-1677, 1990

A>Title: Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional characterization

A:Reference number: A35046; MUID:90175359

A:Accession: A35046

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-610 <HES>

A:Cross-references: GB:M30640; NID:g182047; PIDN:AA52377.1; PID:g182048

R:Bevilacqua, M.P.; Stengelin, S.; Gimbrone Jr., M.A.; Seed, B.
Science 243, 1160-1165, 1989

A>Title: Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils

A:Reference number: A32606; MUID:89162047

A:Accession: A32606

A:Molecule type: mRNA

A:Residues: 1-467, Y, 469-610 <BEV>

A:Cross-references: GB:M4736; NID:g537523; PIDN:AA52376.1; PID:g537524

C:Genetics:

A:Gene: GDB:SELE; ELAM; ESEL; ELAMI

A:Cross-references: GDB:120612; OMIM:131210

A:Map position: 1q22-1q25

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor CKeywords: duplication; glycoprotein; tandem repeat; transmembrane protein

F:1-21/Domain: signal sequence |status predicted <SIG>

F:12-138/Domain: C-type lectin homology <LCH>

F:122-610/Product: endothelial leukocyte adhesion molecule 1 #status predicted <MAT>

F:143-174/Domain: EGF homology <EGF>

F:180-237/Domain: complement factor H repeat homology <FH01>

F:242-299/Domain: complement factor H repeat homology <FH02>

F:304-362/Domain: complement factor H repeat homology <FH03>

F:367-425/Domain: complement factor H repeat homology <FH04>

F:430-488/Domain: complement factor H repeat homology <FH05>

F:493-547/Domain: complement factor H repeat homology <FH06>

F:557-578/Domain: transmembrane #status predicted <TM>

F:25,145,160,179,199,203,265,312,332,503,527/Binding site: carbohydrate (Asn) (covalent),

A:status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1 646 <STR>
A:Cross-references: GB:LI2041; NID:g304246; PIDN:AAA30743.1; PID:g304247
C:Comment: This protein is a Gα2 dependent receptor for myeloid cells.
C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;
C:Keywords: cell adhesion; glycoprotein; phosphatidyl; phospholipid; phosphoprotein; transmembr
E:1-41/Domain: signal sequence #status predicted <Sig>
E:42-646/Product: p-selectin #status predicted <Mat>
F:163-194/Domain: EGF homology <EGF>
F:200-257/Domain: complement factor H repeat homology <FHI>
F:262-319/Domain: complement factor H repeat homology <FH2>
F:324-381/Domain: complement factor H repeat homology <FH3>
F:386-443/Domain: complement factor H repeat homology <FH4>
F:458-515/Domain: complement factor H repeat homology <FH5>
F:520-577/Domain: complement factor H repeat homology <FH6>
F:588-611/Domain: transmembrane #status predicted <TM>
F:612-646/Domain: intracellular #status predicted <Cy>

```

Query Match          39.6%; Score 290; DB 2; Length 646;
Best Local Similarity 39.0%; Pred. No. 2,1e-18;
Matches 48; Conservative 24; Mismatches 51; Indels 0; Gaps 0;

OY      1  CEPLEAPELGMDCDTHPPGNFSSQCAFSCSEGTNLTGIEETTCGPFGNMSSPEPTCOV 60
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      200 CGEFDLPQHVMNCHSLPLGNFSFNHCHGFAEGVALNGPELCLASGLIWTNSPPQV 259
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

OY      61  IQCEPLAPDILGIMNCHPLPLASFSTSACTTCISCBETLIGKKTTCIESSGIMSNPPIC 120
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      260 VOCPALKSPDEGSMSCVQSAEAFQHQSSCSFSCGEGFALVGEVHCTALGVMTAPRVC 319
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

OY      121  QKL 123
      : |
DB      320  KAL 322

RESULT  14
146709
endothelial leukocyte adhesion molecule 1 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 19-May-2000
C:Accession: I46709; I46708
R:Larigan, J.D.; Tsang, T.C.; Rumberger, J.M.; Burns, D.K.
DNA Cell Biol. 11, 149-162, 1992
A:Title: Characterization of cDNA and genomic sequences encoding rabbit ELAM-1: Conse
A:Reference number: I46708; MUID:92189729
A:Accession: I46709
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-551 <LAR1>
A:Cross-references: GB:M01005; NID:g165006; PIDD:AAA31244.1; PID:g165007
A:Accession: I46708
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-307,'T',309-327,'T',329-490,'A',492-551 <LAR2>
A:Cross-references: GB:M01004; NID:g165004; PIDD:AAA31243.1; PID:g165005
C:Genetics:
A:Gene: ELAM1
A:Map position: Iq22-q25
A:Introns: 13/1; I43/1; 179/1; 241/1; 303/1; 366/1; 429/1; 488/1; 525/1; 533/2
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement fa
C:Keywords: duplication; glycoprotein; tandem repeat
F:14-140/Domain: C-type lectin homology <LCB>
F:182-239/Domain: complement factor H repeat homology <FHI>
F:244-301/Domain: complement factor H repeat homology <FHI2>
F:306-364/Domain: complement factor H repeat homology <FHI3>
F:369-427/Domain: complement factor H repeat homology <FHI4>
F:432-486/Domain: complement factor H repeat homology <FHI5>
F:32,45,201,314,321,466/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:32,45,201,314,321,466/Binding site: carbohydrate (Asn) (covalent) #status predicted

```


A:Reference number: A26505; MUID:83048213
 A:Accession: A26505
 A:Molecule type: protein
 A:Residues: 19-20,'Q',22-29,'V',31-33,'Q',35 <SIM>
 R:Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.
 Biochemistry 31, 3626-3634, 1992
 A:Title: Solution structure of the fifth repeat of factor H: A second example of the com
 A:Reference number: A44551; MUID:92232649
 A:Contents: annotation: NMR structure determination, residues 264-292
 R:Kristensen, T.; Wetsel, R.A.; Tack, B.F.
 J. Immunol. 136, 3407-3411, 1986
 A:Title: Structural analysis of human complement protein H: homology with C4b binding pr
 A:Reference number: S10479; MUID:86169701
 A:Accession: S10479
 A:Molecule type: mRNA
 A:Residues: 226-401,'Y',403-449 <KRI>
 A:Cross-references: GB:M12383; NID:g180472; PIDN:AAA52013.1; PID:g180473
 C:Comment: Factor H has also been found bound to cell membranes in an unknown manner. HC
 C:Comment: Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver.
 C:Genetics: <HF1>
 A:Gene: GDB:HF1; HF
 A:Cross-references: GDB:120041; OMIM:134370
 A:Map position: 1932-1932
 C:Genetics: <HF2>
 A:Gene: GDB:HF2; HF
 A:Cross-references: GDB:129095
 A:Map position: 1932-1932
 A:Note: the correspondence between the two loci and the sequences indicated is unclear;
 C:Function:
 A:Description: a cofactor in the inactivation of C3b by serine proteinase I; also increas
 he alternative complement pathway
 A:Pathway: complement alternate pathway
 C:Keywords: alternative splicing; complement alternate pathway; glycoprotein; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-449/Product: complement factor H, short splice form #status experimental <MAT>
 F:21-80/Domain: complement factor H repeat homology <FH01>
 F:85-141/Domain: complement factor H repeat homology <FH02>
 F:146-205/Domain: complement factor H repeat homology <FH03>
 F:210-262/Domain: complement factor H repeat homology <FH04>
 F:246-248/Region: cell attachment (R-G-D) motif
 F:267-320/Domain: complement factor H repeat homology <FH05>
 F:325-385/Domain: complement factor H repeat homology <FH06>
 F:389-442/Domain: complement factor H repeat homology <FH07>
 F:21-66,52-80,85-129,114-141,146-192,178-205,210-251,237-262,267-309,294-320,325-374,357
 F:217/Binding site: carbohydrate (Asn) (covalent) #status absent

A:Title: The complete amino acid sequence of human complement factor H.
 A:Reference number: S00254; MUID:88134059
 A:Accession: S00254
 A:Molecule type: mRNA
 A:Residues: 1-1231 <RIP>
 A:Cross-references: EMBL:Y00716; NID:g31964; PIDN:CNA68704.1; PID:g31965
 A:Note: 402-Tyr was also found
 A:Note: parts of this sequence, including the amino and carboxyl ends of the mature p
 R:Estaller, C.; Schwaebel, W.; Dierich, M.; Weiss, E.H.
 Eur. J. Immunol. 21, 799-802, 1991
 A:Title: Human complement factor H: two factor H proteins are derived from alternativ
 A:Reference number: A60238; MUID:91184292
 A:Accession: A60238
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-56;1177-1231 <EST>
 A:Note: only portions of this 4.3 kilobase mRNA were sequenced
 R:Day, A.J.; Ripchoe, J.; Lyons, A.; McIntosh, B.; Harris, T.J.R.; Sim, R.B.
 Biosci. Rep. 7, 201-207, 1987
 A:Title: Sequence analysis of a cDNA clone encoding the C-terminal end of human compl
 A:Reference number: A54726; MUID:88025472
 A:Accession: A54726
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 'DFRN',579-1231 <DAY>
 A:Cross-references: GB:M17517; NID:g180497; PIDN:AAA52016.1; PID:g180498
 A:Note: parts of this sequence were determined by protein sequencing
 R:Ripchoe, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell, R.D.; Sim, R.B.
 Biosci. Rep. 6, 65-72, 1986
 A:Title: Partial characterization of human complement factor H by protein and cDNA se
 A:Reference number: A61565; MUID:86186123
 A:Accession: A61565
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 'MERCNHLMARK',1050-1057,'T',1059-1102 <RI2>
 R:Sim, R.B.; Discipio, R.G.
 Biochem. J. 205, 285-293, 1982
 A:Title: Purification and structural studies on the complement-system control protein
 A:Reference number: A26505; MUID:83048213
 A:Accession: A26505
 A:Molecule type: protein
 A:Residues: 19-20,'Q',22-29,'V',31-33,'Q',35 <SIM>
 R:Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P
 Biochemistry 31, 3626-3634, 1992
 A:Title: Solution structure of the fifth repeat of factor H: A second example of the
 A:Reference number: A44551; MUID:92232649
 A:Contents: annotation: NMR structure determination, residues 264-292
 R:Kristensen, T.; Wetsel, R.A.; Tack, B.F.
 J. Immunol. 136, 3407-3411, 1986
 A:Title: Cloning of the 1.4-kb mRNA species of human complement factor H reveals a no

Query Match 23.6%; Score 173; DB 1; Length 449;
 Best Local Similarity 28.8%; Pred. No. 4.1e-08;
 Matches 36; Conservative 25; Mismatches 48; Indels 16; Gaps 5;
 QY 14 CTH-----PFGNFS-----FSSQCAFSCEGNTLNG-IEETTCGPRGNSSPPPTQYV 60
 DB 85 CGHGDDPFGFTLLTGNVFEYGVKAVYTCNEGYQLGEIYRCDYDVG-WTNDIPICEV 143
 QY 61 IQCEPLASPDG--IMNSHPIASFSTACTFICSEGTETIGKKTKICSSGSIANSNP 118
 DB 144 VKCLPVPARENGKIVSSAMEPDREYHFGQAVRFVPCNSYKTIKGDEDMHCSDGFEWSEKRP 203
 QY 119 ICOKL 123
 DB 204 KCVEI 208
 RESULT 18
 NBH0H
 complement factor H precursor, long splice form [validated] - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 08-Dec-2000
 C:Accession: S00254; A60238; A54726; A61565; A26505; 172654; S66298
 R:Ripchoe, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.
 Biochem. J. 249, 593-602, 1988

A:Gene: GDB:HFI HF
A:Cross-references: GDB:120041; OMIM:134370
A:Map position: lq32-lq32
C:Genetics: <HF2>
A:Gene: GDB:HF2; HF
A:Cross-references: GDB:129095
A:Map position: lq32-lq32
A>Note: the correspondence between the two loci and the sequences indicated is unclear;
C:Function:
A:Description: a cofactor in the inactivation of C3b by serine proteinase I; also increas
he alternative complement pathway
A:Pathway: complement alternate pathway
C:Superfamily: complement factor H; complement factor H repeat homology
C:Keywords: alternative splicing; complement alternate pathway; glycoprotein; plasma
F:1-18/Domains: signal sequence #status predicted <SIG>
F:19-1229/Product: complement factor H #status experimental <MPY>
F:19-449/Product: complement factor H short splice form #status experimental <MAT>
F:21-80/Domains: complement factor H repeat homology <FH01>
F:25-141/Domains: complement factor H repeat homology <FH02>
F:146-205/Domains: complement factor H repeat homology <FH03>
F:210-262/Domains: complement factor H repeat homology <FH04>
F:246-248/Region: cell attachment (R-G-D) motif
F:325-330/Domains: complement factor H repeat homology <FH05>
F:325-385/Domains: complement factor H repeat homology <FH06>
F:385-442/Domains: complement factor H repeat homology <FH07>
F:448-505/Domains: complement factor H repeat homology <FH08>
F:509-564/Domains: complement factor H repeat homology <FH09>
F:569-623/Domains: complement factor H repeat homology <FH10>
F:630-684/Domains: complement factor H repeat homology <FH11>
F:691-744/Domains: complement factor H repeat homology <FH12>
F:753-803/Domains: complement factor H repeat homology <FH13>
F:811-864/Domains: complement factor H repeat homology <FH14>
F:926-926/Domains: complement factor H repeat homology <FH15>
F:931-984/Domains: complement factor H repeat homology <FH16>
F:989-1043/Domains: complement factor H repeat homology <FH17>
F:1048-1102/Domains: complement factor H repeat homology <FH18>
F:1109-1163/Domains: complement factor H repeat homology <FH19>
F:1167-1228/Domains: complement factor H repeat homology <FH20>
F:21-66-52-80-85-129-114-141-146-192-178-205-210-251-237-262-267-309-294-320-325-374-357
1-803,811-853-839-864,870-915,901-926,931-973,959-984,989-1032,1018-1043,1048-1091,1077
F:529,802,822,882,911/Binding site: carbohydrate (Asn) (covalent) #status absent
F:718,1029,1095/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 23.6%; Score 173; DB 1; Length 1231;
Best Local Similarity 28.8%; Pred. No. 1.1e-07;
Matches 36; Conservative 25; Mismatches 48; Indels 16; Gaps 5;

Oy 14 CTH-----PFGNFS-----FSSQCAFSCSEGTNTG-IETTCGPNGMSPEPTCOV 60
| | | | | :
Db 85 CGHPEDPFEGFFTLGWNVFEGYKAVYTGNBGVQLLEINVRCDTDG-WTNDIPICEV 143
| | | | | :
Oy 61 IQCEPLASPDIG--IMNCSHPLASFPSACTFICSBCTELIGKKTKTCESGSINSNP 118
| | | | | :
Db 144 VKCLFEVTAPENGKIVSSAMEPDREYHFGQAVRFVCNSGYKIEGDPMHCSDGFWSKEP 203
| | | | | :
Oy 119 ICOKL 123
| | | | | :
Db 204 KCVET 208

RESULT 19
136936
Complement receptor 1 - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 02-Jul-1996 #sequence, revision 02-Jul-1996 #text-change 21-Jul-2000
C:Accession: I36937
J:Birmingham, D.J.; Shen, X.P.; Hourcade, D.; Nickells, M.W.; Ackinson, J.P.
J:Immunol. 153, 691-700, 1994
A>Title: Primary sequence of an alternatively spliced form of CRI. Candidate for the 75,
A:Reference number: I36935; MUID:94292799
A:Accession: I36936

```

A>Status: preliminary; translated from GB/EMBL/DBDUT
A:Molecule type: mRNA
A:Residues: 1-2014 <BR1>
A:Cross-references: GB:L24920; NID:g551564; PIDN:AAAS1438.1; PID:g557725
A:Accession: J16937
A>Status: preliminary; translated from GB/EMBL/DBDUT
A:Molecule type: mRNA
A:Residues: 1-397,1751-2014 <BR2>
A:Cross-references: GB:L24921; NID:g557726; PIDN:AAAS1439.1; PID:g557727
C:Genetics:
C:Gene: CRI
C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology
C:Keywords: duplication; glycoprotein; tandem repeat
F:18-74/Domain: complement factor H repeat homology <FH01>
F:79-136/Domain: complement factor H repeat homology <FH02>
F:141-207/Domain: complement factor H repeat homology <FH03>
F:213-268/Domain: complement factor H repeat homology <FH04>
F:222-328/Domain: complement factor H repeat homology <FH05>
F:333-391/Domain: complement factor H repeat homology <FH06>
F:396-462/Domain: complement factor H repeat homology <FH07>
F:470-526/Domain: complement factor H repeat homology <FH08>
F:529-586/Domain: complement factor H repeat homology <FH09>
F:722-778/Domain: complement factor H repeat homology <FH10>
F:1041-1107/Domain: complement factor H repeat homology <FH11>
F:1172-1228/Domain: complement factor H repeat homology <FH12>
F:1233-1291/Domain: complement factor H repeat homology <FH13>
F:1296-1362/Domain: complement factor H repeat homology <FH14>
F:1432-1489/Domain: complement factor H repeat homology <FH15>
F:1635-1681/Domain: complement factor H repeat homology <FH16>
F:1749-1815/Domain: complement factor H repeat homology <FH17>
F:1823-1879/Domain: complement factor H repeat homology <FH18>
F:1884-1940/Domain: complement factor H repeat homology <FH19>

Query Match      22.9%   Score 167.5; DB: 2; Length 2014;
Best Local Similarity    32.6%, Pred.No.5.Se-07;
Matches 46; Conservative 16; Mismatches 56; Indels 23; Gaps 6;

OY 1 CEPLEAPETGTDCVHNPGRGFSFSQCAFCSEGNLTGIERTCGPRGN--WSSPEPT 57
Db 1432 CGPPFPFGNW---HINTDPFGSTVKNCSNEGRLLSGSTICLVSGNVMTDKKAPI 1488
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 58 TCVIDCEP---LSAPDLGIMNCSPPLASFSTSACTFICSECT-----ELIGKKTTIE 108
Db 1489 CEIICECPPTSINSGDF----YSNNRASFGHNQTVVTYCHTGPDEGLFEIVGERSICYT 1544
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 109 SS-----GIMSNSPFCOKLKD 125
Db 1545 SKDDGVGWSSPPRCISTNK 1565
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 20
TJ3012
Complement C3b/C4b receptor, membrane-bound form precursor - human
N:Alternate names: complement C3b/C4b receptor; complement receptor type 1 (CRI); surr
N:Contains: complement C3b/C4b receptor, secreted form
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 24-Nov-1999 #text_change 21-Jul-2000
C:Accession: TJ3012; I56203; A47602; S03291; S03843; A28507; B24748; C24748
R:ViK, D.P.; Wong, W.W.
J: Immunol. 151, 6214-6224, 1993
A>Title: Structure of the gene for the F allele of complement receptor type 1 and seq
A:Reference number: I56203; MUID:94065175
A:Accession: TJ3012
A>Status: preliminary; translated from GB/EMBL/DBDUT
A:Molecule type: DNA
A:Residues: 1-683,'X',685-1133,'X',1135-1471,'X',1473-2489 <VIK1>
A:Cross-references: GB:LI7418; NID:g306678; PIDN:AAB60694.1; PID:g306680
A:Accession: I56203
A>Status: translated from GB/EMBL/DBDUT
A:Molecule type: DNA
A:Residues: 1-683,'X',685-894,'A',896-1000,1451-1471,'X',1473-2489 <VIK2>
A:Cross-references: GB:LI7418; NID:g306678; PIDN:AAB60694.1; PID:g306680

```

R.Wong, W.W.; Cahill, J.M.; Rosen, M.D.; Kennedy, C.A.; Bonaccio, E.T.; Morris, M.J.; Wil
J. Exp. Med. 169, 847-863, 1989
A:Title: Structure of the human C3b gene. Molecular basis of the structural and quantita
A:Reference number: A47602; MUID:89176869
A:Accession: A47602
A:Molecule type: DNA
A:Residues: 1-41 <WON>
A:Cross-references: GB:X14893
R.Hourcade, D.; Miesner, D.R.; Atkinson, J.P.; Holers, V.M.
J. Exp. Med. 168, 1255-1270, 1988
A:Title: Identification of an alternative polyadenylation site in the human C3b/C4b rece
Type 1.
A:Reference number: S03291; MUID:89010527
A:Accession: S03291
A:Molecule type: mRNA
A:Residues: 26-584 <HOU>
A:Cross-references: EMBL:X14362; NID:g30197; PIDN:CA42541.1; PID:g736240
A:Experimental source: Clone CRI-4
R.Klickstein, L.B.; Bartow, T.J.; Miletic, V.; Rabson, L.D.; Smith, J.A.; Fearon, D.T.
J. Exp. Med. 168, 1699-1717, 1988
A:Title: Identification of distinct C3b and C4b recognition sites in the human C3b/C4b r
A:Reference number: S03843; MUID:89035992
A:Accession: S03843
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-894, 'A', 896-1000, 1451-2064, 'T', 2066-2276, 'P', 2278-2299, 'H', 2301-2325, 'T', 2
A:Cross-references: EMBL:Y00816; NID:g30185; PIDN:CA468755.1; PID:g30186
R.Klickstein, L.B.; Wong, W.W.; Smith, J.A.; Weis, J.H.; Wilson, J.G.; Fearon, D.T.
J. Exp. Med. 165, 1095-1112, 1987
A:Title: Human C3b/C4b receptor (CRI). Demonstration of long homologous repeating domain
A:Reference number: A28507; MUID:87168191
A:Accession: A28507
A:Molecule type: mRNA
A:Residues: 953-1221, 'FV', 1224-2064, 'T', 2066-2276, 'P', 2278-2299, 'H', 2301-2325, 'T', 2327-2
A:Cross-references: GB:X05309; NID:g30196; PIDN:CA428933.1; PID:g809019
R.Wong, W.W.; Klickstein, L.B.; Smith, J.A.; Weis, J.H.; Fearon, D.T.
Proc. Natl. Acad. Sci. U.S.A. 82, 7711-7715, 1985
A:Title: Identification of a partial cDNA clone for the human receptor for complement fr
A:Reference number: A94073; MUID:86067975
A:Accession: A24748
A:Molecule type: mRNA
A:Residues: 311-333, 729-745, 831-845 <W02>
A:Cross-references: GB:M11569; NID:g180991; PIDN:AAA52297.1; PID:g180995; GB:M11517; NID
C:Genetics:
A:Gene: GDB:CR1; CD35
A:Cross-references: GDB:119800; OMIM:120620
A:Map position: 1932-1932
A:Introns: 41/1, 101/1, 134/2, 163/1, 296/1, 355/1, 389/2, 418/1, 491/1, 551/1, 584/2, 6
A:1484/2, 1513/1, 1646/1, 1705/1, 1739/2, 1768/1, 1844/1, 1904/1, 1937/2, 1966/1, 2099/1
C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology
C:Keywords: duplication; glycoprotein; transmembrane protein
F:1-41/Domain: signal sequence C3b/C4b receptor, membrane-bound form #status predicted <SIG>
F:42-2489/Product: complement C3b/C4b receptor, secreted form #status predicted <MATS>
F:42-564/Product: complement C3b/C4b receptor, secreted form #status predicted <MATS>
F:43-99/Domain: complement factor H repeat homology <FH01>
F:104-161/Domain: complement factor H repeat homology <FH02>
F:166-232/Domain: complement factor H repeat homology <FH03>
F:238-293/Domain: complement factor H repeat homology <FH04>
F:297-353/Domain: complement factor H repeat homology <FH05>
F:358-416/Domain: complement factor H repeat homology <FH06>
F:421-487/Domain: complement factor H repeat homology <FH07>
F:493-549/Domain: complement factor H repeat homology <FH08>
F:554-611/Domain: complement factor H repeat homology <FH09>
F:616-682/Domain: complement factor H repeat homology <FH10>
F:688-743/Domain: complement factor H repeat homology <FH11>
F:747-803/Domain: complement factor H repeat homology <FH12>
F:808-866/Domain: complement factor H repeat homology <FH13>
F:871-937/Domain: complement factor H repeat homology <FH14>
F:943-999/Domain: complement factor H repeat homology <FH15>
F:1004-1061/Domain: complement factor H repeat homology <FH16>
F:1066-1132/Domain: complement factor H repeat homology <FH17>
F:1138-1193/Domain: complement factor H repeat homology <FH18>
F:1197-1253/Domain: complement factor H repeat homology <FH19>

F:1258-1316/Domain: complement factor H repeat homology <FH20>
F:1321-1387/Domain: complement factor H repeat homology <FH21>
F:1393-1449/Domain: complement factor H repeat homology <FH22>
F:1454-1511/Domain: complement factor H repeat homology <FH23>
F:1516-1582/Domain: complement factor H repeat homology <FH24>
F:1588-1643/Domain: complement factor H repeat homology <FH25>
F:1647-1703/Domain: complement factor H repeat homology <FH26>
F:1708-1766/Domain: complement factor H repeat homology <FH27>
F:1771-1837/Domain: complement factor H repeat homology <FH28>
F:1846-1902/Domain: complement factor H repeat homology <FH29>
F:1907-1964/Domain: complement factor H repeat homology <FH30>
F:1969-2033/Domain: complement factor H repeat homology <FH31>
F:2041-2096/Domain: complement factor H repeat homology <FH32>
F:2100-2156/Domain: complement factor H repeat homology <FH33>
F:2161-2219/Domain: complement factor H repeat homology <FH34>
F:2224-2290/Domain: complement factor H repeat homology <FH35>
F:2298-2354/Domain: complement factor H repeat homology <FH36>
F:2359-2415/Domain: complement factor H repeat homology <FH37>

Query Match 22.3%; Score 163.5; DB 2; Length 2489;
Best Local Similarity 31.9%; Pred. No. 1.5e-06;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;

OY 1 CEPLEAPELGTMDCTHPRGNFSSQCAFSCSESTNLGTIEETCGPGN---WSPEPT 57
Db 1907 CGPPEPFNGMV---HINTDQFGSTVYNSCFRLIGSFTCLVSNVWTKRAPI 1963
OY 58 CQVQCEP---LSAPDGIIMNCSPHSFSTSACTFCSGPT-----ELIGKKTICE 108
Db 1964 CELTSCPEPPIISGDF---YSNNRTSFHNGIVYVTCRTGPGDEQLFELVGRSITCT 2019
OY 109 SS-----GIWNSPSPICQKLDK 125
Db 2020 SKDQYGVWSSPPRCISTNK 2040

RESULT 21
553711
C4BP alpha chain precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Sep-1999 #sequence_revission 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S53711
R:de Frutos, P.G.; Dahlbaeck, B.
Biochim. Biophys. Acta 1261, 285-289, 1995
A:Title: cDNA structure of rabbit C4b-binding protein alpha-chain. Preserved sequence
A:Reference number: S53711; MUID:95226458
A:Accession: S53711
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-597 <DEF>
A:Cross-references: EMBL:235490
C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology
F:50-107/Domain: complement factor H repeat homology <FH1>
F:112-169/Domain: complement factor H repeat homology <FH2>
F:174-234/Domain: complement factor H repeat homology <FH3>
F:239-294/Domain: complement factor H repeat homology <FH4>
F:299-360/Domain: complement factor H repeat homology <FH5>
F:364-422/Domain: complement factor H repeat homology <FH6>
F:426-480/Domain: complement factor H repeat homology <FH7>
F:484-538/Domain: complement factor H repeat homology <FH8>

Query Match 21.8%; Score 160; DB 1; Length 597;
Best Local Similarity 32.5%; Pred. No. 7.9e-07;
Matches 37; Conservative 23; Mismatches 40; Indels 14; Gaps 5;
OY 20 NEFSQQAPSCSGENLGTIEETCG---PFGWSSPEPTCQVQCEPLSAPDLGIMNC 76
Db 128 DFGSQAIEFSCSGYILITSTSHCDIOKGVGWSPLKCEIVKCP---PPN--TING 183
OY 77 SHPLAS---ISFMSACFIFCSGELIGKKKTC---ESSGIWNSPSPICQKLDK 123
Db 128 DFGSQAIEFSCSGYILITSTSHCDIOKGVGWSPLKCEIVKCP---PPN--TING 183

Db 184 KHNGNEDHTGYSSVTVSCNPRFSLGSEASISCFVKNKTVGVWSPSPVCKEI 237

RESULT 22

NEMSH

Complement factor H precursor - mouse

N:Alternate names: protein beta-1-H

C:Species: Mus musculus (house mouse)

C.Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 22-Jun-1999

C.Accession: A26154; 149711; 149728

R.Kristensen, T.; Tack, B.F.

Proc. Natl. Acad. Sci. U.S.A. 83, 3963-3967, 1986

A.Title: Murine protein H is comprised of 20 repeating units, 61 amino acids in length.

A.Reference number: A26154; MUID:86233353

A.Accession: A26154

A.Molecule type: mRNA

A.Residues: 1-1234 <KR1>

A.Cross-references: GB:M12660; NID:9193724; PIDN:AAA37759.1; PID:9387181

R.Natsuume-Sakai, S.; Nonaka, M.; Nonaka, M.; Harada, Y.

J. Immunol. 144, 358-362, 1990

A.Title: Demonstration of an unusual allelic variation of mouse factor H by the complete

A.Reference number: 149711; MUID:90111033

A.Accession: 149711

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-18 <RES>

A.Cross-references: GB:M31979; NID:9193726; PIDN:AAA37762.1; PID:9193729

R.Munoz-Canoves, P.; Tack, B.F.; Vlk, D.P.

Biochemistry 28, 9891-9897, 1989

A.Title: Analysis of complement factor H mRNA expression: Dexamethasone and IFN-gamma in

A.Reference number: 149728; MUID:90148935

A.Accession: 149728

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-19 <RE2>

A.Cross-references: GB:J02891; NID:9193805; PIDN:AAA37795.1; PID:9553926

C.Comment: Two codominant alleles of factor H are present in mice.

C.Comment: Factor H functions as a cofactor in the inactivation of C3b by serine protein

(C5 convertase) in the alternative complement pathway.

C.Genetics:

A.Map position: 1

C.Superfamily: Complement factor H; complement factor H repeat homology

C.Keywords: complement alternate pathway; duplication; glycoprotein; plasma

F.1-18/Domain: signal sequence #status predicted <SIG>

F.19-123/Product: complement factor H #status predicted <MPT>

F.21-80/Domain: complement factor H repeat homology <FH01>

F.85-141/Domain: complement factor H repeat homology <FH02>

F.146-205/Domain: complement factor H repeat homology <FH03>

F.210-262/Domain: complement factor H repeat homology <FH04>

F.246-248/Region: cell attachment (R-G-D) motif

F.267-320/Domain: complement factor H repeat homology <FH05>

F.325-385/Domain: complement factor H repeat homology <FH06>

F.389-442/Domain: complement factor H repeat homology <FH07>

F.448-505/Domain: complement factor H repeat homology <FH08>

F.509-564/Domain: complement factor H repeat homology <FH09>

F.569-622/Domain: complement factor H repeat homology <FH10>

F.620-683/Domain: complement factor H repeat homology <FH11>

F.690-743/Domain: complement factor H repeat homology <FH12>

F.752-802/Domain: complement factor H repeat homology <FH13>

F.808-861/Domain: complement factor H repeat homology <FH14>

F.867-931/Domain: complement factor H repeat homology <FH15>

F.936-989/Domain: complement factor H repeat homology <FH16>

F.994-1048/Domain: complement factor H repeat homology <FH17>

F.1053-1107/Domain: complement factor H repeat homology <FH18>

F.1114-1168/Domain: complement factor H repeat homology <FH19>

F.1172-1233/Domain: complement factor H repeat homology <FH20>

F.21-66,52-80,85-129,114-141,146-192,178-205,210-251,237-262,267-309,294-320,325-374,357

F.802,808-850,836-861,867-920,906-931,936-978,964-989,994-1037,1023-1048,1053-1096,1082-

F.676,721,773,801,1030,1061,1225/Binding site: carbohydrate (asn) (covalent) #status pre

Matches 35; Conservative 25; Mismatches 49; Indels 16; Gaps 5;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

Query Match

Best Local Similarity 21.8%; Score 160; DB 1; Length 1234;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

J. Immunol. 143, 2058-2067, 1989
 A:Title: Murine complement receptor gene family. Identification and characterization
 A:Reference number: A45802; MUID:89381350
 A:Accession: A45802
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 'E', 100-101, 292-961, 964-1025 <KUR>
 A:Cross-references: GB:M9281; NID:q192685; PIDN:AAA37447.1; PID:q387131
 A:Note: the authors failed to translate GGA for residue 421 as G1Y, and CCA for residue 727 as Asn
 C:Superfamily: complement C3b/Epstein-Barr virus receptor; complement factor H repeat
 C:Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane
 F:11-11/Domain: signal sequence #status predicted <SIG>
 F:12-1025/Product: complement C3d/Epstein-Barr virus receptor 2 (15-repeat form) #status
 F:12-973/Domain: extracellular #status predicted <EXT>
 F:14-73/Domain: complement factor H repeat homology <FH01>
 F:82-138/Domain: complement factor H repeat homology <FH02>
 F:146-202/Domain: complement factor H repeat homology <FH03>
 F:207-263/Domain: complement factor H repeat homology <FH04>
 F:268-334/Domain: complement factor H repeat homology <FH05>
 F:342-398/Domain: complement factor H repeat homology <FH06>
 F:402-458/Domain: complement factor H repeat homology <FH07>
 F:463-514/Domain: complement factor H repeat homology <FH08>
 F:519-585/Domain: complement factor H repeat homology <FH09>
 F:594-649/Domain: complement factor H repeat homology <FH10>
 F:654-704/Domain: complement factor H repeat homology <FH12>
 F:709-769/Domain: complement factor H repeat homology <FH13>
 F:778-833/Domain: complement factor H repeat homology <FH14>
 F:841-897/Domain: complement factor H repeat homology <FH15>
 F:907-958/Domain: complement factor H repeat homology <FH16>
 F:968-999/Domain: transmembrane #status predicted <TM>
 F:990-1025/Domain: intracellular #status predicted <INT>

Query Match 20.5%; Score 150.5; DB 1; Length 1025;
 Best Local Similarity 26.3%; Pred. No 9.4e-06;
 Matches 52; Conservative 12; Mismatches 55; Indels 79; Gaps 8;

QY 1 CE-----PLEAPELGTMDCTHPFG---NFSFSQCAFSCSEGTNLGLIETTCGPGFMSS 53
 Db 138 CESDFLEPCSPSLPTIHNHHTGQHVDFVAGLSVYSCPEGLTLGKTKICLSSGMDWG 197
 QY 54 PEPCQVQVQ-----EPLS----- 67
 Db 198 VIPTCKEAQCHPGKFKPGQVKEPLSLQVGTYYFSCNEGTYOLOGQPSQCVIEQKAIW 257
 QY 68 -----APDLGINCSHPPLASFSS---FTSACFFTCSESTE-----LIGKKRTI 106
 Db 258 TKRPCKREILCPRPVYVANGSH-TGSPFSENPVPGSTVYTCDPDPKGVSPFTLLIGERTIN 316
 QY 107 C-----ESSGIMSNPSPTC 120
 Db 317 CTTGSGQKTGIMSGAPVYC 334

RESULT 26
 NBHHC4
 C4b-binding protein alpha chain precursor - human
 N:Alternate names: C4BP; proline-rich protein
 C:Species: Homo sapiens (man)
 C:Date: 13-Aug-1986 #sequence, revision 30-Jun-1993 #text, change 22-Jun-1999
 C:Accession: A33568; S02372; A00326; A24182; A93134; S29492; A31785; I52244; A03210
 R:Matsumuchi, T.; Okemura, S.; Aso, T.; Sata, T.; Niho, Y.
 Biochem. Biophys. Res. Commun. 165, 138-144, 1989
 A:Title: Molecular cloning of the cDNA coding for proline-rich protein (PRP): identitl
 A:Reference number: A33568; MUID:90073699
 A:Accession: A33568
 A:Molecule type: mRNA
 A:Residues: 1-597 <MA1>
 A:Cross-references: GB:M31452; NID:q190501; PIDN:AAA36507.1; PID:q190502
 A:Note: the authors translated the codon GGA for residue 492 as Glu
 A:Lintin, S.J.; Lewin, A.R.; Reid, K.B.M.
 FEBS Lett. 232, 328-332, 1988

A:Title: Derivation of the sequence of the signal peptide in human C4b-binding protein A.
A:Reference number: S02372; MUID:86242821
A:Accession: 502372
A:Molecule type: mRNA
A:Residues: 17-81 <LI2>
A:Cross-references: EMBL:X07853
A:Note: although the sequence determined extends to residue 9 above, these authors assigned
R:Chung, L.P.; Bentley, D.R.; Reid, K.B.M.
Biochem. J. 230, 133-141, 1985
A:Title: Molecular cloning and characterization of the cDNA coding for C4b-binding protein A.
A:Reference number: A90326; MUID:86025405
A:Accession: A90326
A:Molecule type: mRNA
A:Residues: 80-597 <CH2>
A:Cross-references: GB:X02865; NID:g29564; PIDN:CAA26617.1; PID:g29565
A:Note: 92-Thr and 357-His were also found
R:Linlin, S.J.; Reid, K.B.M.
FEBS Lett. 204, 77-81, 1986
A:Title: Studies on the structure of the human C4b-binding protein gene.
A:Reference number: A24182; MUID:86301119
A:Accession: A24182
A:Molecule type: DNA
A:Residues: 203-288 <LIN>
A:Cross-references: EMBL:X04284; EMBL:X04296
R:Rodriguez de Cordoba, S.; Sanchez-Corral, P.; Rey-Campos, J.
Exp. Med. 173, 1073-1082, 1991
A:Title: Structure of the gene coding for the alpha polypeptide chain of the human complement C4.
A:Reference number: A43023; MUID:9121619
A:Contents: annotation; exon-intron boundaries
R:Chung, L.P.; Gagnon, J.; Reid, K.B.M.
Mol. Immunol. 22, 427-435, 1985
A:Title: Amino acid sequence studies of human C4b-binding protein: N-terminal sequence determined by cyanogen bromide treatment.
A:Reference number: A93134; MUID:85296001
A:Accession: A93134
A:Molecule type: protein
A:Residues: 49-81 <CH1>
A:Note: this paper reports amino-terminal sequences of the intact protein and of a number of fragments.
R:Hessling, M.; Katers, D.; Takeya, H.; van't Veer, C.; Hackeng, T.M.; Iwanaga, S.; Bouma
FEBS Lett. 317, 228-232, 1993
A:Title: The region Ser(333)-Arg(356) of the alpha-chain of human C4b-binding protein is a reference number: S29492; MUID:93146164
A:Reference number: S29492; MUID:93146164
A:Accession: S29492
A>Status: preliminary
A:Molecule type: protein
A:Residues: 381-404 <HES>
R:Suzuki, K.; Nishioke, J.
J. Biol. Chem. 263, 17034-17039, 1988
A:Title: Binding site for vitamin K-dependent protein S on complement C4b-binding protein A.
A:Reference number: A31785; MUID:89034204
A:Accession: A31785
A:Molecule type: protein
A:Residues: 495-505, 'X', 507-510, 'X', 512-515 <SU2>
A:Note: this peptide appears to bind protein S
R:Dahlback, B.; Smith, C.A.; Muller-Eberhard, H.J.
Proc. Natl. Acad. Sci. U.S.A. 80, 3461-3465, 1983
A:Title: Visualization of human C4b-binding protein and its complexes with vitamin K-dependent protein S.
A:Reference number: A9350; MUID:83221615
A:Contents: annotation; electron microscopy; three-dimensional structure; ligand binding
R:Asao, T.; Okumura, S.; Matsuguchi, T.; Sakamoto, N.; Sata, T.; Niho, Y.
Biochem. Biophys. Res. Commun. 174, 222-227, 1991
A:Title: Genomic organization of the alpha chain of the human C4b-binding protein gene.
A:Reference number: I52244; MUID:91131199
A:Accession: I52244
A>Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-597 <ASO>
A:Cross-references: GB:M62486; NID:g190498; PIDN:AAA36506.1; PID:g190500
A:Comment: C4b controls the classical pathway of complement activation. It binds as a component of the C4b2a complex (C3 convertase) by dissociating the complement fragment C4b2a. C4b occurs in plasma in two forms, both of which bind complement fragment C4b2a. The beta chain binds the vitamin K-dependent plasma protein S. A minor form is called C4b2a-1.
A:Comment: C4b occurs in plasma in two forms, both of which bind complement fragment C4b2a. The beta chain binds the vitamin K-dependent plasma protein S. A minor form is called C4b2a-1.

C:Comment: The molecule has a central body supporting seven tentacles (alpha chains)
C:Genetics:
A:Gene: GDB:C4BPA
A:Cross-references: GDB:120568; OMIM:120830
A:Map position: 1q32-1q32
A:Introns: 48/1; 110/1; 143/2; 172/1; 236/1; 297/1; 362/1; 425/1; 482/1; 540/3
C:Complex: octamer of seven alpha chains and one beta chain
C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology
C:Keywords: acute phase; chylomicron; complement pathway; duplication; glycoprotein;
F:1148/Domains: signal sequence #status predicted <Sig>
F:46-597/Product: C4b-binding protein alpha chain #status predicted <Mat>
F:50-108/Domains: complement factor H repeat homology <FH1>
F:113-170/Domains: complement factor H repeat homology <FH2>
F:175-234/Domains: complement factor H repeat homology <FH3>
F:235-294/Domains: complement factor H repeat homology <FH4>
F:295-360/Domains: complement factor H repeat homology <FH5>
F:364-422/Domains: complement factor H repeat homology <FH6>
F:381-404/Region: complement C4b binding #status predicted
F:426-480/Domains: complement factor H repeat homology <FH7>
F:484-538/Domains: complement factor H repeat homology <FH8>
F:221, 506, 528/Binding site: carbohydrate (Asn) (covalent) #status experimental

```

Query Match          20.4%; Score 149.5; DB 1; Length 597;
Best Local Similarity 29.1%; Pred. No. 6,8e-06;
Matches 37; Conservative 18; Mismatches 53; Indels 19; Gaps 4;

Oy 14 CTHP-----TGNFSSQCAFCSCSEGTNLGIETTCGPRC---NWSSPEPTQY 60
      |||      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 CHHPELRNGOYEIKTIDLSFGSIOEFSCSEGFLLIGSTTSRCEYODRGVGHNPJPOCB 172
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 61 IOCEPLSAPDLICIMNCSPHLASFSTSACTFICSECTELIGKKTKIC---SSSGTMSNP 116
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 VACKP--PPDINGNRHSGEENFYATGFSVTYSCDPRFSLGHASISCTYENETIGVWRPS 230
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 117 SPICOKL 123
      ||| |||
Db 231 PPTCEKI 237

RESULT 27
B45900
complement C3d/Epstein-Barr virus receptor precursor alternative splice form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C:Accession: B45900; C45900
R:Kurtz, C.B.; O'Toole, E.; Christensen, S.M.; Weis, J.H.
J. Immunol. 144, 3581-3591, 1990
A:Title: The murine complement receptor gene family. IV. Alternative splicing of Cr2
A:Reference number: A45900; MUID:90229754
A:Accession: B45900
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-363 <KUR>
A:Accession: C45900
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-19, 'E', '84-363 <KUZ>
A:Residues: 1-19, 'E', '84-363 <KUZ>
C:Superfamily: complement C3d/Epstein-Barr virus receptor; complement factor H repeat
F:23-81/Domain: complement factor H repeat homology <FH1>
F:86-145/Domain: complement factor H repeat homology <FH2>
F:154-210/Domain: complement factor H repeat homology <FH3>
F:218-274/Domain: complement factor H repeat homology <FH4>
F:279-335/Domain: complement factor H repeat homology <FH5>

Query Match          20.2%; Score 148; DB 2; Length 363;
Best Local Similarity 32.1%; Pred. No. 5,7e-06;
Matches 43; Conservative 17; Mismatches 58; Indels 16; Gaps 7;

Oy 1 CB---PLEAPELGTMDCTHPG-----NFSFSCAFSCSEGTNLGIETTCGPRGNSS 53
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 CESDPELEPSPPTJHTNGHTGOHVDQFVAGLSVYVSCSPGYLLTKGKTKICLSSDDMG 269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 10 GNGEYGAKKVYTCDEGVQWGMNFRECDTNG--WTNDIPICEVKKCLPVTPEBNGKIFS 68
 Oy 75 NCSHPLASFSTTSACTFICSEGTGELIGKKTKTCESGSIWNSPSPICQKL 123
 Db 69 DALEPDEYTYGVAVQFECNSGYMLDGPQIHCISAGVWSAETPKCWEI 117

RESULT 31

T20239
 hypothetical protein C54G4.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T20239
 R:Millinson, J.
 submitted to the EMBL Data Library, June 1996
 A:Reference number: 219241
 A:Accession: T20239
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-868 <M18>
 A:Cross-references: EMBL:Z75533; PIDN:CAA99822.1; GSPDB:GN00019; CESP:C54G4.4
 A:Experimental source: clone C54G4
 C:Genetics:
 A:Gene: CESP:C54G4.4
 A:Map position: 1
 A:Introns: 189/2; 254/2; 291/3; 339/1; 380/2; 536/1; 562/3; 661/1; 756/3; 829/1

Query Match 20.1%; Score 147.5; DB 2; Length 868;
 Best Local Similarity 31.4%; Pred. No. 1.5e-05;
 Matches 33; Conservative 14; Mismatches 47; Indels 11; Gaps 3;

Oy 23 FSSQCAFSCSEGNITGIEET-----CGPFGNWSPEPTCOYIQCEPLSAPDLGIMN 75
 Db 559 FESEVEYECNSGMHLANSPSYSRLRRVCSQSDGIMSGSEPTCELVDG---GRPL-IAN 614
 Oy 76 CSHPLASFSTTSACTFICSEGTGELIGKKTKTCESGSIWNSPSPIC 120
 Db 615 GRVDESSTFESANVTCHQGRILGPESLMCGDRGEMQPATPFC 659

RESULT 32

T34288
 hypothetical protein F47C12.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34288
 R:Murray, J.; Mohlmann, P.; Sansone, J.
 submitted to the EMBL Data Library, June 1996
 A:Description: The sequence of C. elegans cosmid F47C12.
 A:Reference number: 221499
 A:Accession: T34288
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1827 <MUR>
 A:Cross-references: EMBL:U61946; PIDN:AA24388.1; GSPDB:GN00022; CESP:F47C12.1
 A:Experimental source: Strain Bristol N2; clone F47C12
 C:Genetics:
 A:Gene: CESP:F47C12.1
 A:Map position: 4
 A:Introns: 66/2; 221/1; 282/1; 404/1; 500/1; 649/3; 750/1; 997/2; 1152/3; 1279/1; 1346/1

Query Match 20.1%; Score 147.5; DB 2; Length 1827;
 Best Local Similarity 27.1%; Pred. No. 3.1e-05;
 Matches 39; Conservative 18; Mismatches 48; Indels 39; Gaps 7;

Oy 1 CEPL---EAPL---GTMDCTHPRFNGFSFSSQCAFSCSBS---TNTLGIEETTCGPPGNKS 52
 Db 280 COPICGAVPOJANGFVDSA---TNVSTGQVKKSCHKGFFPAGSKDVETVYCGEIGKWK 336
 Oy 53 SPEPTCOYIQCEPLSAPDLGIMNSHPLASF-----FTSACFTICSEGTCL 99

Db 337 IP-PACKATCO-----PLAQFSGNDRRLFEGDGTGYGVSFRDCHSGYRR 381
 Oy 100 IGKKTICSSSGSIWNSPSPICQKL 123
 Db 382 EGVESSLCKSDGTWSSKQPNCTKV 405

RESULT 33

S57953
 C4BP protein alpha chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Aug-1999
 C:Accession: S57953
 R:Hillarp, A.; Thern, A.; Dahlback, B.
 submitted to the EMBL Data Library, July 1995
 A:Description: Molecular cloning of rat C4b-binding protein alpha- and beta-chains:
 A:Reference number: S57953
 A:Accession: S57953
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-558 <H1>
 A:Cross-references: EMBL:Z50051; NID:9899379; PIDN:CAA90391.1; PID:9899380
 C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology
 F:15-72/Domain: complement factor H repeat homology <FH1>
 F:77-134/Domain: complement factor H repeat homology <FH2>
 F:139-199/Domain: complement factor H repeat homology <FH3>
 F:204-258/Domain: complement factor H repeat homology <FH4>
 F:263-324/Domain: complement factor H repeat homology <FH5>
 F:328-386/Domain: complement factor H repeat homology <FH6>
 F:390-443/Domain: complement factor H repeat homology <FH7>
 F:447-501/Domain: complement factor H repeat homology <FH8>

Query Match 19.9%; Score 146; DB 2; Length 558;
 Best Local Similarity 28.6%; Pred. No. 1.3e-05;
 Matches 42; Conservative 20; Mismatches 57; Indels 28; Gaps 7;

Oy 1 CEPLAPELG-----TMDCTHPRG-----NFSFSCQCAFSCSEGNITGIEETNC 45
 Db 60 CKPLKQWQININACVKKSCNPGDILONGKVEYKIDFLFGSQIEFSCSEGYILGSSSTVCE 119
 Oy 46 --GPFGNWSPEPTCOYIQCEPLSAPDLGIMNSH---PLASFSTTSACTFICSEGTGELI 100
 Db 120 IOGKGVSMDSPLPECVIAKCG--MPD--ISNGKINGREEFYTRSSVTVYACDDFTLL 175
 Oy 101 GKKRTIC---ESSGIWNSPSPICQKL 123
 Db 176 GNASITCTVYNKTVGVWSPSPPTCERI 202

RESULT 34

I46001
 C4b-binding protein alpha chain - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: I46001; S43190
 R:Hillarp, A.; Thern, A.; Dahlback, B.
 J. Immunol. 153, 4190-4199, 1994
 A:Title: Bovine C4b binding protein. Molecular cloning of the alpha- and beta-chains
 A:Reference number: I46001; MUID:95015909
 A:Accession: I46001
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-610 <H1>
 A:Cross-references: EMBL:Z31693; NID:9469117; PIDN:CAA83498.1; PID:9469118
 C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology
 F:50-107/Domain: complement factor H repeat homology <FH1>
 F:112-169/Domain: complement factor H repeat homology <FH2>
 F:174-234/Domain: complement factor H repeat homology <FH3>
 F:239-294/Domain: complement factor H repeat homology <FH4>
 F:299-362/Domain: complement factor H repeat homology <FH5>
 F:366-425/Domain: complement factor H repeat homology <FH6>

C:Keywords: chylomicron; duplication; glycoprotein; HDL; heparin binding; lipid binding;
F:1-19/Domain: signal sequence (fragment) #status predicted <SIG>
F:20-345/Product: apolipoprotein H #status experimental <MAT>
F:23-79/Domain: complement factor H repeat homology <FH1>
F:84-137/Domain: complement factor H repeat homology <FH2>
F:142-200/Domain: complement factor H repeat homology <FH3>
F:205-260/Domain: complement factor H repeat homology <FH4>
F:264-325/Domain: complement factor H repeat homology <FH5>
F:23-66, 51-79, 84-124, 110-137, 142-188, 174-200, 205-248, 234-260, 264-315, 300-325, 307-345/DLS
F:92,162,193,253/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 19.3%; Score 141.5; DB 1; Length 345;
Best Local Similarity 31.5%; Pred. No. 2,1e-05;
Matches 39; Conservative 15; Mismatches 63; Indels 7; Gaps 6;

Oy 1 CEPLEAPELGTMDCTHPR-GNEFS-SSQCAFSCSEGTNLGTIEETTCGPRGNSPEPTC 58
Db 142 CPPPTIPFASISLVKPLAGNNSFYGSKAVFCLPHHMFQNDVTCTEHGMMTQ-LDEC 200
Oy 59 QVIOCEPLSPADLTGIMNSHPL-ASFSTSACTFICSEGTGLGKKTKICSSGIMSNPS 117
Db 201 REVRCPSPSPDNGFVN--HPANPLYKYKDTATFCGHEFTYSLDGEVEVCSKFGMWS-AQ 257
Oy 118 PRQ 121
Db 258 PSCK 261

RESULT 38
PL0009
Complement C3d/Epstein-Barr virus receptor precursor - human
N:Alternate names: complement receptor 2; CR2/CD21
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence revision 07-Jul-1995 #text-change 22-Jun-1999
C:Accession: J10028; A39958; A32036; A24319; B24319; D24319; E24319; F24319; PLD
R:Wels, J.D.; Toohtaker, L.E.; Smith, J.A.; Wels, J.H.; Fearon, D.T.
J. Exp. Med. 167, 1047-1066, 1988
A:Title: Structure of the human B lymphocyte receptor for C3d and the Epstein-Barr virus
A:Reference number: J10028; MUID:86171282
A:Accession: J10028
A:Molecule type: mRNA
A:Residues: 1-1091 <WEI>
A:Note: nucleotides 1566-1625 are missing from figure 1; therefore, residues 522-542 have
R:Moore, M.D.; Cooper, N.R.; Tack, B.F.; Nemerow, G.R.
Proc. Natl. Acad. Sci. U.S.A. 84, 9194-9198, 1987
A:Title: Molecular cloning of the cDNA encoding the Epstein-Barr virus/C3d receptor (com
A:Reference number: A39958; MUID:88097454
A:Accession: A39958
A:Molecule type: mRNA
A:Residues: 1-456, 'G', 457-658, 'R', 646-669, 'R', 671-816, 'NCSAEVLKMWIERAR', 835-840, 'L', 84
A:Cross-references: GB:J03565; MID:9181919; PID:AAA35784.1; PID:9181920
R:Fujioka, A.; Harley, J.B.; Frank, M.B.; Gruner, B.A.; Frazier, B.; Holers, V.M.
J. Biol. Chem. 264, 2118-2125, 1989
A:Title: Genomic organization and polymorphisms of the human C3d/Epstein-Barr virus rece
A:Reference number: A32036; MUID:89123277
A:Accession: A32036
A:Molecule type: mRNA
A:Residues: 1-456, 'G', 457-658, 'R', 646-669, 'R', 671-816, 'NCSAEVLKMWIERAR', 835-840, 'L', 84
A:Cross-references: GB:J04463
R:Wels, J.D.; Fearon, D.T.; Klickstein, L.B.; Wong, W.W.; Richards, S.A.; De Bruyn Kops,
Proc. Natl. Acad. Sci. U.S.A. 83, 5639-5643, 1986
A:Title: Identification of a partial cDNA clone for the C3d/Epstein-Barr virus receptor
of complement.
A:Reference number: A94114; MUID:86287311
A:Accession: A24319
A:Molecule type: protein
A:Residues: 226-230, 'XIIQ', 257-267, 332-341, 583-591, 'Q', 593, 'D', 595-596, 728-735 <WE2>
A:Experimental source: B-lymphoblastoid cell lines SB and Raji
C:Genetics: GDB:CR2
A:Cross-references: GDB:119802; OMIM:120650
A:Map position: 1q32-1q32

C:Superfamily: complement C3d/Epstein-Barr virus receptor; complement factor H repeat
C:Keywords: alternative splicing; duplication; glycoprotein; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-1091/Product: complement receptor 2 (16-repeat form) #status predicted <MAT>
F:21-658, 718-1091/Product: complement receptor 2 (15-repeat form) #status predicted <MAT>
F:23-82/Domain: complement factor H repeat homology <FH1>
F:91-146/Domain: complement factor H repeat homology <FH2>
F:154-210/Domain: complement factor H repeat homology <FH3>
F:276-342/Domain: complement factor H repeat homology <FH4>
F:351-406/Domain: complement factor H repeat homology <FH5>
F:470-465/Domain: complement factor H repeat homology <FH6>
F:470-521/Domain: complement factor H repeat homology <FH7>
F:526-592/Domain: complement factor H repeat homology <FH8>
F:601-656/Domain: complement factor H repeat homology <FH9>
F:660-716/Domain: complement factor H repeat homology <FH10>
F:720-772/Domain: complement factor H repeat homology <FH11>
F:777-837/Domain: complement factor H repeat homology <FH12>
F:846-901/Domain: complement factor H repeat homology <FH13>
F:909-965/Domain: complement factor H repeat homology <FH14>
F:970-1026/Domain: complement factor H repeat homology <FH15>
F:1034-1056/Domain: transmembrane #status predicted <TM>
F:1057-1091/Domain: intracellular #status predicted <INT>
F:121,127,294,372,622,698,858,881,919/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 19.2%; Score 141; DB 1; Length 1091;
Best Local Similarity 29.4%; Pred. No. 7,1e-05;
Matches 35; Conservative 16; Mismatches 62; Indels 6; Gaps 3;

Oy 7 PELGTMDCTHPRGN--FSFSSQCAFSCSEGTNLGTIEETTCGPRGNSPEPTC 63
Db 911 PPKTPNGNHTGNTARRSPGMSILYSCDQGLLVGEALLDCTHGTMSQAPRKEVNC 970
Oy 64 EPLSPADLTGIMNSHPLASFSTSACTFICSEGTGLGKKTKICSSGIMSNPSPIQ 121
Db 971 S--SPADMGIQIGLEPRKMYGYAVLTLECDGYMLEGSPQSQSDHWNPIAVCR 1027

RESULT 39
BBH0
Complement factor B precursor [validated] - human
N:Alternate names: C3 convertase; C3 proactivator; glycine-rich beta-glycoprotein; he
N:Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) Bb fragmen
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1984 #sequence revision 05-Aug-1994 #text-change 08-Dec-2000
C:Accession: S34075; A44622; A19188; A19947; B19947; B25971; S14339; A44628;
R:Mejia, J.E.; Jahn, I.; de la Salle, H.; Hauptmann, G.
submitted to the EMBL Data Library, March 1993
A:Reference number: S34075
A:Accession: S34075
A:Molecule type: mRNA
A:Residues: 1-764 <MEU>
A:Cross-references: EMBL:X72875; MID:q297568; PID:CAA51389.1; PID:q297569
A:Note: the authors translated the codon TAC at 519 as Thr; the nucleic acid translat
R:Moore, M.D.; Cooper, N.R.; Tack, B.F.; Nemerow, G.R.
Proc. Natl. Acad. Sci. U.S.A. 79, 5661-5665, 1982
A:Title: Isolation of cDNA clones for the human complement protein factor B, a class
A:Reference number: A44622; MUID:83039428
A:Accession: A44622
A:Molecule type: mRNA
A:Residues: 467-546, 550-595, 752-764 <MOO>
A:Cross-references: GB:J00186
A:Note: the authors translated the codon TAC at 519 as Thr; the nucleic acid translat
R:Moore, M.D.; Cooper, N.R.; Tack, B.F.; Nemerow, G.R.
Proc. Natl. Acad. Sci. U.S.A. 79, 5661-5665, 1982
A:Title: Isolation of cDNA clones for the human complement protein factor B, a class
A:Reference number: A20751; MUID:84161997
A:Accession: A20751
A:Molecule type: protein; mRNA
A:Residues: 26-764 <MO>
A:Cross-references: GB:K01566
A:Note: nucleic acid translation differs from the sequence shown in having 300-Leu, 3
A:Note: 736-Ser was also found

A>Note: glycosylation sites were determined
 R:Christie, D.L.; Gagnon, J.
 Biochem. J. 209, 61-70, 1983
 A:Title: Amino acid sequence of the Bb fragment from complement factor B. Sequence of th
 A:Reference number: A19188; M0ID:83204002
 A:Contents: the final paper in a series documenting the sequence, glycosylation site, an
 A:Accession: A19188
 A:Molecule type: protein
 A:Residues: 260-296, 'T', 298-764 <CHR>
 R:Campbell, R.D.; Porter, R.R.
 Proc. Natl. Acad. Sci. U.S.A. 80, 4464-4468, 1983
 A:Title: Molecular cloning and characterization of the gene coding for human complement
 A:Reference number: A19947; M0ID:83273641
 A:Accession: A19947
 A:Molecule type: DNA
 A:Residues: 346-764 <CAM>
 A:Cross-references: GB:J00125
 A:Accession: B19947
 A:Molecule type: mRNA
 A:Residues: 339-509 <CAI>
 A:Cross-references: GB:J00126; NID:9187723; PIDN:AAA36226.1; PID:9553536
 R:Wu, L.; Morley, B.J.; Campbell, R.D.
 Cell 48, 331-342, 1987
 A:Title: Cell-specific expression of the human complement protein factor B gene: evidenc
 A:Reference number: A25971; M0ID:87102880
 A:Accession: B25971
 A:Molecule type: DNA
 A:Residues: 1-99 <MUI>
 A:Cross-references: GB:M15082; NID:9187699; PIDN:AAA59625.1; PID:9553534
 R:Niemann, W.A.; Bhown, A.S.; Miller, E.J.
 Biochem. J. 274, 473-480, 1991
 A:Title: The principal site of glycation of human complement Factor B.
 A:Reference number: S14339; M0ID:91174758
 A:Accession: S14339
 A:Molecule type: protein
 A:Residues: 270-329 <NIE>
 A>Note: binding site for carbohydrate to lysine under artificial conditions
 R:Morley, B.J.; Campbell, R.D.
 EMBO J. 3, 153-157, 1984
 A:Title: Internal homologies of the Ba fragment from human complement component factor H
 A:Reference number: A44628; M0ID:84158524
 A:Accession: A44628
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 16-225, 'F', 227-259 <MOR>
 R:Schwaebel, W.; Lutting, B.; Sokolowski, T.; Estaller, C.; Weiss, E.H.; Meyer zum Busche
 Immunobiology 180, 221-232, 1993
 A:Title: Human complement factor B: CDNA cloning, nucleotide sequencing, phenotypic cony
 A:Reference number: I54409; M0ID:94041399
 A:Accession: I54409
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1764 <RES>
 A:Cross-references: GB:S67310; NID:9452937; PIDN:AAI13989.1; PID:94261689
 R:Horluch, T.; Kim, S.; Matsumoto, M.; Watanabe, I.; Fujita, S.; Volanakis, J.E.
 Mol. Immunol. 30, 1587-1592, 1993
 A:Title: Human complement factor B: CDNA cloning, nucleotide sequencing, phenotypic cony
 A:Reference number: I57824; M0ID:94067177
 A:Accession: I57824
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-31, 'Q', 33-764 <RE2>
 A:Cross-references: GB:I15702; NID:9291921; PIDN:AAA16820.1; PID:9291922
 C:Comment: 292-Cys has a free sulhydryl.
 C:Genetics:
 A:Gene: GDB:BF
 A:Cross-references: GDB:I19726; OMIM:138470
 A:Map position: 6p21.3-6p21.3
 A:Introns: 21/3; 99/3; 346/1; 390/1; 424/1; 470/1; 502/3; 542/1; 593/2; 619/1; 652/3; 69
 A>Note: the list of introns may be incomplete
 A>Note: gene is located in the major histocompatibility complex, class III region
 C:Complex: complement factor B initially forms an inactive complex with complement facto
 ment factor C3b forming active C3/C5 convertase; Ba is released

C:Function:
 A:Description: Bb is a serine proteinase; C3/C5 convertase cleaves complement C3 alph
 a1
 A:Pathway: complement alternate pathway
 C:Superfamily: complement C2; complement factor H repeat homology; trypsin homology;
 C:Keywords: acute phase; complement alternate pathway; duplication; glycoprotein; hyd
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-764/Product: complement factor B #status experimental <MAT>
 F:26-259/Product: complement factor Ba fragment #status experimental <BAR>
 F:37-98/Domain: complement factor H repeat homology <PH1>
 F:103-158/Domain: complement factor H repeat homology <PH2>
 F:165-218/Domain: complement factor H repeat homology <PH3>
 F:260-764/Product: C3/C5 convertase Bb fragment #status experimental <BBF>
 F:268-458/Domain: von Willebrand factor type A repeat homology <VFA>
 F:482-752/Domain: trypsin homology #status atypical <TRY>
 F:37-76,62-98,103-145,131-158,165-205,191-218,478-596,511-527,599-615,656-682,695-725
 F:122,142,285,378/Binding site: carbohydrate (asn) (covalent) #status experimental
 F:259-260/Cleavage site: Arg-Lys (complement factor D) #status experimental
 F:526,576,699/Active site: His, Asp, Ser #status experimental

Query Match 19.2% Score 140.5; DB 1; Length 764;
 Best Local Similarity 25.2%; Pred. No. 5.5e-05;
 Matches 35; Conservative 14; Mismatches 55; Indels 35; Gaps 4;

Qy 1 CEPLAEPLGLTMDCTHPFGN-----FSSQCAFSCSEGTNLGIEETTCGPGMW 51
 Db 98 CRAITHCP-----PHDFENGGEVWPSPYNNVSDSEIFHCYDGLTGLSANRRCQVNGRW 151
 Qy 52 SSPEPTCOVIOCEPLSPADGLGIMNCSP-----LASFFTSACTFCSEGTGLIK 102
 Db 152 SGQPAIC-----DNGAGCYCNPGLPIGTRKYGSQRLSDSVTHCSRLTLRGS 200
 Qy 103 KRTICSSGIMSNPSPIQ 121
 Db 201 QRRTQSGSGSNSTPSCQ 219

RESULT 40
 WNVZSP
 apolipoprotein H homolog precursor - vaccinia virus
 N:Alternate names: 35K secretory protein; C3L protein; virokin
 C:Species: vaccinia virus
 A>Note: host Homo sapiens (man)
 C:Date: 31-Dec-1989 #sequence_revision 30-Jun-1990 #text_change 22-Jun-1999
 C:Accession: A31005; B42504
 R:Kotwal, G.J.; Moss, B.
 Nature 335, 176-178, 1988
 A:Title: Vaccinia virus encodes a secretory polypeptide structurally related to compl
 A:Reference number: A31005; M0ID:88318974
 A:Accession: A31005
 A:Molecule type: DNA
 A:Residues: 1-263 <ROT>
 A:Cross-references: GB:X13166; NID:960690; PIDN:CAA31564.1; PID:960691
 A:Experimental sources: strain WR
 R:Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
 Virology 179, 517-563, 1990
 A:Title: Appendix to "The complete DNA sequence of vaccinia virus".
 A:Reference number: A42501
 A:Accession: A42501
 A:Molecule type: DNA
 A:Residues: 1-263 <GOE>
 A:Cross-references: GB:M35027; NID:9335317; PIDN:AAA47997.1; PID:9335345
 A:Experimental sources: strain Copenhagen
 R:Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
 Virology 179, 247-266, 1990
 A:Title: The complete DNA sequence of vaccinia virus.
 A:Reference number: A42531; M0ID:91021027
 A:Contents: annotation: possible protein-coding frames
 A>Note: neither amino acid nor nucleotide sequence is given
 C:Superfamily: herpesvirus complement control protein; complement factor H repeat hom
 C:Keywords: duplication; extracellular protein
 F:1-19/Domain: signal sequence #status predicted <SIG>

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:22:09 : Search time 85.16 Seconds
(without alignments)
392.017 Million cell updates/sec

Title: US-09-119-209-2_COPY_197_328
Perfect score: 733
Sequence: 1 CEPLAEPGLTMDCTHPEFGN.....MSNPSPICQKLDKSFMSIKE 132

Scoring table:
BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 869228 seqs, 252910224 residues

Total number of hits satisfying chosen parameters: 869228

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents AA-New:*
1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	733	100.0	372	US-09-791-537-152667	Sequence 152667, Ap
2	727	99.2	184	US-10-212-054-1437	Sequence 1437, Ap
3	727	99.2	372	US-09-791-537-22816	Sequence 22816, A
4	727	99.2	372	US-09-791-537-42657	Sequence 42657, A
5	727	99.2	372	US-09-791-537-84593	Sequence 84593, A
6	727	99.2	385	US-09-791-537-51391	Sequence 51391, A
7	724	98.8	372	US-09-791-537-42659	Sequence 42659, A
8	719	98.1	363	US-09-791-537-58446	Sequence 58446, A
9	719	98.1	385	US-09-791-537-113060	Sequence 113060, A
10	713	97.3	372	US-09-791-537-69658	Sequence 69658, A
11	713	97.3	385	US-09-791-537-53844	Sequence 53844, A
12	690	94.1	372	US-09-791-537-42655	Sequence 42655, A
13	690	94.1	372	US-09-791-537-42658	Sequence 42658, A
14	674	92.0	341	US-10-211-364-1171	Sequence 1171, Ap
15	674	92.0	341	US-10-212-054-1128	Sequence 1128, Ap
16	674	92.0	341	US-10-212-778-1158	Sequence 1158, Ap
17	599	81.7	376	US-09-791-537-132144	Sequence 132144, A
18	537	73.3	370	US-09-791-537-50403	Sequence 50403, A
19	535	73.0	372	US-09-791-537-81233	Sequence 81233, A
20	528	72.0	372	US-09-791-537-6693	Sequence 6693, Ap
21	518	70.7	323	US-09-791-537-53485	Sequence 53485, A
22	518	70.7	360	US-09-791-537-60503	Sequence 60503, A
23	518	70.7	372	US-09-791-537-37750	Sequence 37750, A
24	318	43.4	646	US-09-791-537-18300	Sequence 18300, A
25	315	43.0	740	US-09-791-537-32024	Sequence 32024, A
26	315	43.0	830	US-09-791-537-22819	Sequence 22819, A

27	315	43.0	830	5	US-09-791-537-35618	Sequence 35618, A
28	313	42.7	768	5	US-09-791-537-86038	Sequence 86038, A
29	309	42.2	616	5	US-09-791-537-33025	Sequence 33025, A
30	307	41.9	740	5	US-09-791-537-32023	Sequence 32023, A
31	306	41.7	754	5	US-09-791-537-14651	Sequence 14651, A
32	305.5	41.7	612	5	US-09-791-537-73416	Sequence 73416, A
33	305.5	41.7	618	5	US-09-791-537-81892	Sequence 81892, A
34	305.5	41.7	619	5	US-09-791-537-37749	Sequence 37749, A
35	300	40.9	611	5	US-09-791-537-118878	Sequence 118878, A
36	300	40.9	768	5	US-09-791-537-20989	Sequence 20989, A
37	300	40.9	768	5	US-09-791-537-37753	Sequence 37753, A
38	300	40.9	549	5	US-09-791-537-50409	Sequence 50409, A
39	299.5	40.9	549	5	US-09-791-537-50406	Sequence 50406, A
40	296	40.4	610	1	PCR-US02-23913-357	Sequence 357, App
41	296	40.4	610	5	US-09-791-537-22813	Sequence 22813, A
42	296	40.4	610	5	US-09-791-537-121834	Sequence 121834, A
43	296	40.4	610	6	US-10-205-823-357	Sequence 357, App
44	291	39.7	649	5	US-09-791-537-93873	Sequence 93873, A
45	290	39.6	646	5	US-09-791-537-84829	Sequence 84829, A

ALIGNMENTS

```
RESULT 1
US-09-791-537-152667
Sequence 152667, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biocomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 152667
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (11)..(11)
OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-152667

Query Match          100.0%; Score 733; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.56-57;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEPLAEPGLTMDCTHPEFGNFSFSSQACSCSGTTLTGIEFTGCPFGMNSPEPTCOV 60
DB 197 CEPLAEPGLTMDCTHPEFGNFSFSSQACSCSGTTLTGIEFTGCPFGMNSPEPTCOV 256
QY 61 IOCEPLAEPDLGIMNCSHPLASFSTACTFTCSGTELDIGKKTICSSGIMNSPPTC 120
DB 257 IOCEPLAEPDLGIMNCSHPLASFSTACTFTCSGTELDIGKKTICSSGIMNSPPTC 316
QY 121 OKLDKSFMSIKE 132
DB 317 OKLDKSFMSIKE 328

RESULT 2
US-10-212-054-1437
Sequence 1437, Application US/10212054
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P02120CIN
```

```
; CURRENT APPLICATION NUMBER: US/10/212,054
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2164
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1437
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-212-054-1437
```

```
Query Match          99.2%; Score 727; DB 6; Length 184;
Best Local Similarity 99.2%; Pred. No. 4.1e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 9 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 68
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 120
    |||||||
DB 69 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 128
QY 121 OKLDKSFMSIKE 132
    |||||||
DB 129 OKLDKSFMSIKE 140
```

```
RESULT 3
US-09-791-537-22816
; Sequence 22816, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22816
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-22816
```

```
Query Match          99.2%; Score 727; DB 5; Length 372;
Best Local Similarity 99.2%; Pred. No. 8.6e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 197 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 120
    |||||||
DB 257 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 316
QY 121 OKLDKSFMSIKE 132
    |||||||
DB 317 OKLDKSFMSIKE 328
```

```
RESULT 4
US-09-791-537-42657
; Sequence 42657, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
```

```
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42657
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pan troglodytes
US-09-791-537-42657
```

```
Query Match          99.2%; Score 727; DB 5; Length 372;
Best Local Similarity 99.2%; Pred. No. 8.6e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 197 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 120
    |||||||
DB 257 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 316
QY 121 OKLDKSFMSIKE 132
    |||||||
DB 317 OKLDKSFMSIKE 328
```

```
RESULT 5
US-09-791-537-84593
; Sequence 84593, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84593
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-84593
```

```
Query Match          99.2%; Score 727; DB 5; Length 372;
Best Local Similarity 99.2%; Pred. No. 8.6e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 197 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGTIEETTCGPGNMSSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 120
    |||||||
DB 257 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTELIGKKTKICSSGIMNSPSPIC 316
QY 121 OKLDKSFMSIKE 132
    |||||||
DB 317 OKLDKSFMSIKE 328
```

```
RESULT 6
US-09-791-537-51391
; Sequence 51391, Application US/09791537
```

```

: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ. ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 51391
: LENGTH: 385
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-791-537-51391

```

Query Match	99.28;	Score 727;	DB 5;	Length 385;
Best Local Similarity	99.28;	Pred. No. 8.9e-57;		
Matches 131; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	CEPLEAEALGIMCTHFGNFSFSSQCAFSCSESTNLGTIEETCGPFGMWSPEPFCQY	60
Db	210	CEPLEAEALGIMCTHPLGNFSFSQCAFSCSESTNLGTIEETCGPFGMWSPEPFCQY	269
QY	61	IQCEPLAPDLGIMNCNCHPLASFSTACTPICSEGTLELIGKKTKTICSSGIMNSPDC	120
Db	270	IQCEPLAPDLGIMNCNCHPLASFSTACTPICSEGTLELIGKKTKTICSSGIMNSPDC	329
QY	121	OKLDSFSMIKE	132
Db	330	OKLDSFSMIKE	341

```

RESULT 7
US-09-791-537-42659
: Sequence 42659, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Blonmlix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 42659
: LENGTH: 372
: TYPE: PRT
: ORGANISM: Pongo pygmaeus
: US-09-791-537-42659

```

Query Match	98.8%;	Score 724;	DB 5;	Length 372;
Best Local Similarity	98.5%;	Pred. No. 1.6e-56;		
Matches 130;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Oy	1	CEPLEAPELGIMDCNHPGCFNSFSQCAFCSCEGNTLGIETTCGFCFGMSSPEPTCOV	60	
Db	197	CEPLEAPELGIMDCNHPGCFNSFSQCAFCSCEGNTLGIETTCGFCFGMSSPEPTCOV	256	
Oy	61	IQCEPLSPADGIGMNCSHPLASFSTSACTFCSEGTETLIGKKTICSSGIMSNPSPIC	120	
Db	257	IQCEPLSPADGIGMNCSHPLASFSTSACTFCSEGTETLIGKKTICSSGIMSNPSPIC	316	
Oy	121	OKLDKSFMSIKE	132	
Db	317	OKLDKSFMSIKE	328	

```

RESULT      8
US-09-791-537-58446
: Sequence 58446, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomolx, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 58446
: LENGTH: 363
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-791-537-58446

```

Query Match	98.1%	Score 719;	DB 5;	Length 363;
Best Local Similarity	98.5%;	Pred. No. 4.3e-56;		
Matches 130;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

[illegible]

```

RESULT      9
US-09-791-537-113060
: Sequence 113060, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biocomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 113060
: LENGTH: 385
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-791-537-113060

```

Query Match	98.1%	Score 719	DB 5	Length 385
Best Local Similarity	98.5%	Pred. No. 4.6e-56		
Matches 130	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1	CEPLEAPELGMDCHPGRGFSFSQCAFSCSEGTNLGIEFTTGGPGGMMSSPEPTCOV	60	
Db	210	CEPLEAPELGMDCHSLGNGFSFSQCAFSCSEGTNLGIEFTTGGPGGMMSSPEPTCOV	269	
QY	61	IQCEPLAPDGLGIMMCHSHPLASFSTSACTFICSGTELIGKKKTIICSSSGIWMSPDIC	120	
Db	270	IQCEPLAPDGLGIMMCHSHPLASFSTSACTFICSGTELIGKKKTIICSSSGIWMSPDIC	329	
QY	121	OKLDKSFMSIKE	132	

Db 330 OKDKSFSMIKE 341

RESULT 10

US-09-791-537-69658

; Sequence 69658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 69658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-69658

Query Match 97.3%; Score 713; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 1.5e-55;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 197 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 256
QY 61 IQCEPLAPDLGINCNSHPLASFSTACTFICSEGTLLIGKKTICSSGIMNSPSPIC 120
Db 257 IQCEPLAPDLGINCNSHPLASFSTACTFICSEGTLLIGKKTICSSGIMNSPSPIC 316
QY 121 OKDKSFSMIKE 132
Db 317 OKDKSFSMIKE 328

RESULT 11

US-09-791-537-53844

; Sequence 53844, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 53844
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53844

Query Match 97.3%; Score 713; DB 5; Length 385;
Best Local Similarity 97.0%; Pred. No. 1.6e-55;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 210 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 269
QY 61 IQCEPLAPDLGINCNSHPLASFSTACTFICSEGTLLIGKKTICSSGIMNSPSPIC 120
Db 270 IQCEPLAPDLGINCNSHPLASFSTACTFICSEGTLLIGKKTICSSGIMNSPSPIC 329

QY 121 OKDKSFSMIKE 132
Db 330 OKDKSFSMIKE 341

RESULT 12

US-09-791-537-42655

; Sequence 42655, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 42655
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-791-537-42655

Query Match 94.1%; Score 690; DB 5; Length 372;
Best Local Similarity 92.4%; Pred. No. 1.7e-53;
Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 197 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 256
QY 61 IQCEPLAPDLGINCNSHPLASFSTACTFICSEGTLLIGKKTICSSGIMNSPSPIC 120
Db 257 IQCEPLAPDLGINCNSHPLASFSTACTFICSEGTLLIGKKTICSSGIMNSPSPIC 316
QY 121 OKDKSFSMIKE 132
Db 317 OKDKSFSMIKE 328

RESULT 13

US-09-791-537-42658

; Sequence 42658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 42658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Papio hamadryas
US-09-791-537-42658

Query Match 94.1%; Score 690; DB 5; Length 372;
Best Local Similarity 92.4%; Pred. No. 1.7e-53;
Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 197 CEPLAPDLGMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 256

QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMSNPSPIC 120
|||||
Db 257 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMSNPSPIC 316
QY 121 OKDKSFSMIKE 132
|||||
Db 317 OKDKSFSMIKE 328

RESULT 14

US-10-211-364-1171
; Sequence 1171, Application US/10211364
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P216CIN
; CURRENT APPLICATION NUMBER: US/10/211,364
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: 09/760,486
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1778
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1171
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-211-364-1171

Query Match 92.0%; Score 674; DB 6; Length 341;
Best Local Similarity 99.2%; Pred. No. 4,1e-52;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPFNFSFSSQCAFCSEGTNLGTGIEETTCGPFGNWSSPEPTCOY 60
|||||
Db 219 CEPLAPELGTMDCTHPFNFSFSSQCAFCSEGTNLGTGIEETTCGPFGNWSSPEPTCOY 278

QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMSNPSPIC 120
|||||
Db 279 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMSNPSPIC 338

QY 121 Q 121
Db 339 Q 339

RESULT 15
US-10-212-054-1328
; Sequence 1328, Application US/10212054
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P212CIN
; CURRENT APPLICATION NUMBER: US/10/212,054
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2164
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1328
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-054-1328

Query Match 92.0%; Score 674; DB 6; Length 341;
Best Local Similarity 99.2%; Pred. No. 4,1e-52;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPFNFSFSSQCAFCSEGTNLGTGIEETTCGPFGNWSSPEPTCOY 60
|||||
Db 219 CEPLAPELGTMDCTHPFNFSFSSQCAFCSEGTNLGTGIEETTCGPFGNWSSPEPTCOY 278

QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMSNPSPIC 120
|||||
Db 279 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMSNPSPIC 338

QY 121 Q 121
Db 339 Q 339

RESULT 16

US-10-212-778-1158
; Sequence 1158, Application US/10212778
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM026CIN
; CURRENT APPLICATION NUMBER: US/10/212,778
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/758,449
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1478
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1158
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-778-1158

Query Match 92.0%; Score 674; DB 6; Length 341;
Best Local Similarity 99.2%; Pred. No. 4,1e-52;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPFNFSFSSQCAFCSEGTNLGTGIEETTCGPFGNWSSPEPTCOY 60
|||||
Db 219 CEPLAPELGTMDCTHPFNFSFSSQCAFCSEGTNLGTGIEETTCGPFGNWSSPEPTCOY 278

QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMSNPSPIC 120

Db 279 IOCEPLAPDLGIMNCSHPLASFSFTSACTFICSEGTLLGKKTICSSGIMNSNPIC 338
QY 121 Q 121
Db 339 Q 339

RESULT 17
US-09-791-537-132144

; Sequence 132144, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 132144

; LENGTH: 376

; TYPE: PRT

; ORGANISM: Oryctolagus cuniculus

US-09-791-537-132144

Query Match 81.7%; Score 599; DB 5; Length 376;
Best Local Similarity 80.3%; Pred. No. 2,1e-45;
Matches 106; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 60
Db 197 CEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 256
QY 61 IOCEPLAPDLGIMNCSHPLASFSFTSACTFICSEGTLLGKKTICSSGIMNSNPIC 120
Db 257 IOCEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 316
QY 121 OKLDRFSMIKE 132
Db 317 OKLDRFSMIKE 328

RESULT 18

US-09-791-537-50403

; Sequence 50403, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 50403

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Bos taurus

US-09-791-537-50403

Query Match 73.3%; Score 537; DB 5; Length 370;
Best Local Similarity 68.9%; Pred. No. 6.7e-40;
Matches 91; Conservative 19; Mismatches 22; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 60
Db 197 CEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 256
QY 61 IOCEPLAPDLGIMNCSHPLASFSFTSACTFICSEGTLLGKKTICSSGIMNSNPIC 120
Db 257 IOCEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 316
QY 121 OKLDRFSMIKE 132
Db 317 OKLDRFSMIKE 328

Db 197 CVPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 256
QY 61 IOCEPLAPDLGIMNCSHPLASFSFTSACTFICSEGTLLGKKTICSSGIMNSNPIC 120
Db 257 IOCEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 316
QY 121 OKLDRFSMIKE 132
Db 317 OKLDRFSMIKE 328

RESULT 19
US-09-791-537-81233

; Sequence 81233, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 81233

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-791-537-81233

Query Match 73.0%; Score 535; DB 5; Length 372;
Best Local Similarity 71.2%; Pred. No. 1e-39;
Matches 94; Conservative 15; Mismatches 23; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 60
Db 197 CEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 256
QY 61 IOCEPLAPDLGIMNCSHPLASFSFTSACTFICSEGTLLGKKTICSSGIMNSNPIC 120
Db 257 IOCEPLAPDLGIMTCTHFGNFSFSQCAFSCSGTNTLGIEFTTCGPGNWSPEPTCOV 316
QY 121 OKLDRFSMIKE 132
Db 317 OKLDRFSMIKE 328

RESULT 20

US-09-791-537-6693

; Sequence 6693, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6693

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Rattus sp

US-09-791-537-6693

Query Match 72.0%; Score 528; DB 5; Length 372;
Best Local Similarity 70.5%; Pred. No. 4.2e-39;
Matches 93; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

Best Local Similarity 42.1%; Pred. No. 3.5e-20;
Matches 51; Conservative 22; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60
    |||||:| ||||| |||:| ||||| |||:| ||||| |||:| ||||| |||:| |||||
Db 324 CEPLSPVHGSMDCSPSLRAFYDTCNSFRCAEGFMLRGADIVRCDNLGQWTAPAPVC 383
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKTKICSSGIMNSPPTC 120
    |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db 384 LOCODLPAPEKQVNCSPHPCAFRYOSTCSTFCDEGSSLVGASVLCLETGNMSAPAPEC 443
OY 121 Q 121
    |
Db 444 Q 444
```

RESULT 25
US-09-791-537-32024
Sequence 32024, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 32024

LENGTH: 740

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-32024

Query Match 43.0%; Score 315; DB 5; Length 740;

Best Local Similarity 43.9%; Pred. No. 7.5e-20;

Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60
    |||||:| ||||| |||:| ||||| |||:| ||||| |||:| ||||| |||:| |||||
Db 262 COHLEAPSEGTMDCVHPLTAFAYGSSCKFECPGYRVRGIDMLRCIDSGHMSAPLPTCEA 321
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKTKICSSGIMNSPPTC 120
    |||||:| ||||| |||:| ||||| |||:| ||||| |||:| ||||| |||:| |||||
Db 322 ISCEPLSPVHGSMDCSPSLRAFYDTCNSFRCAEGFMLRGADIVRCDNLGQWTAPAPVC 381
OY 121 Q 123
    |
Db 382 QAL 384
```

RESULT 26
US-09-791-537-22819
Sequence 22819, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 22819

LENGTH: 830

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-22819

Query Match 43.0%; Score 315; DB 5; Length 830;

Best Local Similarity 43.9%; Pred. No. 8.5e-20;

Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60
    |||||:| ||||| |||:| ||||| |||:| ||||| |||:| ||||| |||:| |||||
Db 324 COHLEAPSEGTMDCVHPLTAFAYGSSCKFECPGYRVRGIDMLRCIDSGHMSAPLPTCEA 383
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKTKICSSGIMNSPPTC 120
    |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db 384 ISCEPLSPVHGSMDCSPSLRAFYDTCNSFRCAEGFMLRGADIVRCDNLGQWTAPAPVC 443
OY 121 Q 123
    |
Db 444 QAL 446
```

RESULT 27
US-09-791-537-35618
Sequence 35618, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 35618

LENGTH: 830

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-35618

Query Match 43.0%; Score 315; DB 5; Length 830;

Best Local Similarity 43.9%; Pred. No. 8.5e-20;

Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60
    |||||:| ||||| |||:| ||||| |||:| ||||| |||:| ||||| |||:| |||||
Db 324 COHLEAPSEGTMDCVHPLTAFAYGSSCKFECPGYRVRGIDMLRCIDSGHMSAPLPTCEA 383
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKTKICSSGIMNSPPTC 120
    |||||:| ||||| |||:| ||||| |||:| ||||| |||:| ||||| |||:| |||||
Db 384 ISCEPLSPVHGSMDCSPSLRAFYDTCNSFRCAEGFMLRGADIVRCDNLGQWTAPAPVC 443
OY 121 Q 123
    |
Db 444 QAL 446
```

RESULT 28
US-09-791-537-86038
Sequence 86038, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 86038

LENGTH: 768

```
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-86038
```

```
Query Match          42.7%; Score 313; DB 5; Length 766;
Best Local Similarity 43.9%; Pred. No. 1.2e-19;
Matches 54; Conservative 17; Mismatches 52; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGIEETTCGPGFMWSSPEPTCOV 60
    | : | | | | | | | | | | : | : | : | : | : | : | : | : | : | : |
DB 262 CGSELEPQHVLMNCSPHLGNFSFNSQCSFHCIDGYVNGPSKLECLASGIWTKKPPQCLA 321
OY 61 IOCEPLSAPDLGIMNCSPHLASFSFTSACTFTCSCTELIGKKTKICSSGIMNSPSTIC 120
    | : | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 322 IACEPLSPLHSGMDCFPSTGANGVNSCTCFRCTEGFVLMGNDIHCADLGOWTAPAPVC 381
OY 121 QKL 123
    | : |
DB 382 EAL 384
```

```
RESULT 29
US-09-791-537-32025
```

```
; Sequence 32025, Application US/09791537
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
```

```
; APPLICANT: Debe, Derek
```

```
; APPLICANT: Danzer, Joseph
```

```
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
```

```
; TITLE OF INVENTION: METHODS OF USE THEREOF
```

```
; FILE REFERENCE: 261/210
```

```
; CURRENT APPLICATION NUMBER: US/09/791,537
```

```
; CURRENT FILING DATE: 2001-02-22
```

```
; NUMBER OF SEQ ID NOS: 153055
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 32025
```

```
; LENGTH: 616
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
US-09-791-537-32025
```

```
Query Match          42.2%; Score 309; DB 5; Length 616;
Best Local Similarity 41.5%; Pred. No. 2.1e-19;
Matches 51; Conservative 23; Mismatches 49; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGIEETTCGPGFMWSSPEPTCOV 60
    | : | | | | | | | | | | : | : | : | : | : | : | : | : | : | : |
DB 200 CGSELEPQHVLMNCSPHLGNFSFNSQCSFHCIDGYVNGPSKLECLASGIWTKKPPQCLA 259
OY 61 IOCEPLSAPDLGIMNCSPHLASFSFTSACTFTCSCTELIGKKTKICSSGIMNSPSTIC 120
    | : | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 260 ACCPPLKIPERGMMTCCLHSAKAFOHSSCSFCEGFAVGPEVVOCTASGWTAPAPVC 319
OY 121 QKL 123
    | : |
DB 320 KAL 322
```

```
RESULT 30
US-09-791-537-32023
```

```
; Sequence 32023, Application US/09791537
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
```

```
; APPLICANT: Debe, Derek
```

```
; APPLICANT: Danzer, Joseph
```

```
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
```

```
; TITLE OF INVENTION: METHODS OF USE THEREOF
```

```
; FILE REFERENCE: 261/210
```

```
; CURRENT APPLICATION NUMBER: US/09/791,537
```

```
; CURRENT FILING DATE: 2001-02-22
```

```
; NUMBER OF SEQ ID NOS: 153055
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 32023
```

```
; LENGTH: 740
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
US-09-791-537-32023
```

```
Query Match          41.9%; Score 307; DB 5; Length 740;
Best Local Similarity 40.7%; Pred. No. 3.9e-19;
Matches 50; Conservative 24; Mismatches 49; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGIEETTCGPGFMWSSPEPTCOV 60
    | : | | | | | | | | | | : | : | : | : | : | : | : | : | : | : |
DB 200 CGSELEPQHVLMNCSPHLGNFSFNSQCSFHCIDGYVNGPSKLECLASGIWTKKPPQCLA 259
OY 61 IOCEPLSAPDLGIMNCSPHLASFSFTSACTFTCSCTELIGKKTKICSSGIMNSPSTIC 120
    | : | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 260 ACCPPLKIPERGMMTCCLHSAKAFOHSSCSFCEGFAVGPEVVOCTASGWTAPAPVC 319
OY 121 QKL 123
    | : |
DB 320 KAL 322
```

```
RESULT 31
```

```
US-09-791-537-14651
```

```
; Sequence 14651, Application US/09791537
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
```

```
; APPLICANT: Debe, Derek
```

```
; APPLICANT: Danzer, Joseph
```

```
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
```

```
; TITLE OF INVENTION: METHODS OF USE THEREOF
```

```
; FILE REFERENCE: 261/210
```

```
; CURRENT APPLICATION NUMBER: US/09/791,537
```

```
; CURRENT FILING DATE: 2001-02-22
```

```
; NUMBER OF SEQ ID NOS: 153055
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 14651
```

```
; LENGTH: 754
```

```
; TYPE: PRT
```

```
; ORGANISM: Canis familiaris
```

```
US-09-791-537-14651
```

```
Query Match          41.7%; Score 306; DB 5; Length 754;
Best Local Similarity 40.5%; Pred. No. 4.8e-19;
Matches 49; Conservative 24; Mismatches 48; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGIEETTCGPGFMWSSPEPTCOV 60
    | : | | | | | | | | | | : | : | : | : | : | : | : | : | : | : |
DB 185 CGDFDLPQHVLMNCSPHLGNFSFNSQCSFHCIDGYVNGPSKLECLASGIWTKKPPQCLA 244
OY 61 IOCEPLSAPDLGIMNCSPHLASFSFTSACTFTCSCTELIGKKTKICSSGIMNSPSTIC 120
    | : | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 245 TOCPPLKIPERGMMTCCLHSAKAFOHSSCSFCEGFAVGPEVVOCTASGWTAPAPVC 304
OY 121 Q 121
    | : |
DB 305 E 305
```

```
RESULT 32
```

```
US-09-791-537-73416
```

```
; Sequence 73416, Application US/09791537
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
```

```
; APPLICANT: Debe, Derek
```

```
; APPLICANT: Danzer, Joseph
```

```
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
```

```
; TITLE OF INVENTION: METHODS OF USE THEREOF
```

```
; FILE REFERENCE: 261/210
```



```

: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 20989
: LENGTH: 768
: TYPE: PR1
: ORGANISM: Mus musculus
: US-09-791-537-20989

```

Query Match	40.9%	Score 300	DB 5	Length 768
Best Local Similarity	42.3%	Pred. No.	1.7e-18	
Matches	52	Conservative	17	Mismatches 54
				Indels 0
				Gaps 0

QY	1	C E P L E A R E L G M D C T H F P G N S F S S Q C A F S O S B E G T N I L G I E R T T C G F F G M W S S E P F T C Q Y	60
		: : : : : : : : : : : : : : : : :	
D b	200	G K G V N I D Q H V L M N C S H P L G E F S S Q C T F S C A E E Y E I D G G E L O C L A S G I W T N N P P C D A	259
QY	61	I Q C E P L A P D L G I M N C S H P L A S F S F T S A C T F I C S E G T E L L G K K T T E S S G I M S N S P I C	120
		: : : : : : : : : : : : : : : : :	
D b	260	V O C S L E A P P H G T M A C M H P I A F A Y D S S C K F E C P O G Y R A R O S N T L H C T G S G O W S E P L P T C	319
QY	121	O K L 123	
	:	:	
D b	320	E A I 322	

```

RESULT 37
US-09-791-537-37753
: Sequence 37753, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 37753
: LENGTH: 768
: TYPE: PRF
: ORGANISM: Mus musculus
: US-09-791-537-37753

```

Query Match	40.9%	Score 300;	DB 5;	Length 768;
Best Local Similarity	42.3%	Pred. No. 1.7e-18;		
Matches 52% Conservative	17;	Mismatches 34;	Indels 0;	Gaps 0;

```

QY      1 CEPLEADELGTMOCTHFGNNSFSSCOAFSCSEGTNLGTIEHTTCGPFNMSSPEPTCY 60
      2  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      200 CGKNIHQHVLNCSHPLGSEFSFNSQCTFSCAEEYELDGPGEIQLCLASGIWTNNPKCDA 259
      61  IOCEPLSAPDLGIMNCSHPLASFSTACTFTICSEGTLELKKKTIICSSGIMSNPSPIC 120
      260 VQCSLEAPPHGIMACHMPLIAFAFYDSSCKFECPGGRANGSNTLHCTGSGQWSEPLPTC 319
QY      121  QKL 123
      320  :  :
      320  EAI 322

```

RESULT 38

```

US-09-791-537-50409
: Sequence 50409, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 50409
: LENGTH: 769
: TYPE: PRT
: ORGANISM: Ovis aries
US-09-791-537-50409

```

Query Match	40.9%	Score 300:	DB 5:	Length 769:
Best Local Similarity	39.8%	Pred No. 1.7e-18:		
Matches 49; Conservative	24;	Mismatches 50;	Indels 0;	Gaps 0

OY 1 CEPLEPAPELGNIDCHTPGPNSSFSOCAFSCSEGNLTGIETTCGPGNMSSEPICOV 60
 :
Db 200 CGEFDLPQNVHNMNCSPHPGNSFKSQCSCEHCABGYALNGPRELBCLASGITNTNSPQCYA 255

OY 61 IOCEPLSAPDLCIMMCSHPILASFSTACTCTICGEHTLLKKTITCESSIMSNPSPIC 120
 :
Db 260 VCCPALKPSEGSMSCFISAKAFORHQSCSCSCCEGLTVPBHVCHTALGVATNPAPVC 319

```

RESULT 39
US-09-791-537-50406
: Sequence 50406, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 50406
: LENGTH: 549
: TYPE: PRT
: ORGANISM: Rattus norvegicus
US-09-791-537-50406

```

Query Match	40.98;	Score 299.5;	DB 5,	Length 549;
Best Local Similarity	41.08;	Pred. No. 1.3e-18;		
Matches 50;	Conservative 23;	Mismatches 48;	Indels 1;	Gaps 1;

0Y 1 CEELEAPLGLMCHHPGNGFSFSSCOAFSCSEBETNLGTEET -GGPFGWSSPEPTCQ 59
180 CQEGEYPHGSLNCHTFFGLFSTYSSCSFSEBERGVSSMTTYKRCISSEGMSPAPACH 239
0Y 60 VIQCEPLASPDGLMCHSHPLASFSETSACTFTCSBETELIGKKKTCISSGAIWSNPDI 119
240 VEECALQPHAGVKKCSNDSGSPWMTTCFTDEEGRRVAGANLDCSTSSGVNDIKPS 299

Db 300 CK 301

```
RESULT 40
PCT-US02-23913-357
; Sequence 357, Application PC/TUS0223913
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-04PC
; CURRENT APPLICATION NUMBER: PCT/US02/23913
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-23913-357
```

```
Query Match 40.4%; Score 296; DB 1; Length 610;
Best Local Similarity 41.3%; Pred. No. 3e-18;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY 1 CEPLAPELLGTMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPPGMNSPEPTQYV 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 CTALSPESHSLVCSHPLGNFSYSSCSISCDRGLPSSMETWQCMSSGEMSAPIPCNV 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 IQCEPLASAPDLGIMNCHPLASFSTACTFTICSEGTLLIGKKTICSSGSTMSPIC 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 VECDAVTNPANGFVECFQNPQSGFPMNTTCTFDCEGFEMLGAQSLQCTSSGNWDNEKPTC 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 Q 121
Db 300 K 300
```

Search completed: September 7, 2002, 10:22:11
Job time: 559 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:20:38 : Search time 304.64 Seconds
(without alignments)
152.513 Million cell updates/sec

Title: US-09-119-209-2_COPY_197_328

Perfect score: 733
Sequence: 1 CEPLEAPELCTMDCTHFGN.....WSNPSPICKDKSFSMIKE 132

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCRTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	733	100.0	372	15	US-09-119-209-2
2	727	99.2	184	21	US-09-760-443-1437
3	727	99.2	184	21	US-09-760-475-3252
4	727	99.2	372	1	PCT-US01-26675-3
5	727	99.2	382	1	US-09-760-475-2123
6	727	99.2	1078	26	US-60-207-315-428
7	727	99.2	1078	26	US-60-212-659-523

8	727	99.2	1078	26	US-60-230-435-1751	Sequence 1751, Ap
9	713	97.3	385	1	PCT-US92-03970-2	Sequence 2, App11
10	713	97.3	385	1	PCT-US94-00909-2	Sequence 2, App11
11	713	97.3	385	4	US-08-008-459-2	Sequence 2, App11
12	713	97.3	385	7	US-08-340-539-2	Sequence 2, App11
13	713	97.3	385	8	US-08-410-569-2	Sequence 2, App11
14	674	92.0	341	21	US-09-758-449-1158	Sequence 1158, Ap
15	674	92.0	341	21	US-09-760-443-1328	Sequence 1328, Ap
16	518	70.7	372	15	US-09-119-209-4	Sequence 4, App11
17	315	43.0	830	1	PCT-US94-09395-4	Sequence 4, App11
18	315	43.0	830	8	US-08-449-687B-4	Sequence 4, App11
19	315	43.0	830	24	US-10-020-141-10	Sequence 10, App1
20	306	41.7	700	26	US-60-207-315-467	Sequence 467, App
21	296	40.4	610	1	PCT-US99-28965-19	Sequence 19, App1
22	296	40.4	610	10	US-08-657-753-2	Sequence 2, App11
23	296	40.4	610	11	US-08-770-435-3	Sequence 3, App11
24	296	40.4	610	16	US-09-266-091A-2	Sequence 2, App11
25	296	40.4	610	16	US-09-266-091A-2	Sequence 2, App11
26	296	40.4	610	21	US-09-784-356-122	Sequence 122, App
27	296	40.4	610	22	US-09-802-640-36	Sequence 36, App1
28	296	40.4	610	22	US-09-857-670-19	Sequence 19, App1
29	296	40.4	610	24	US-10-021-660-122	Sequence 122, App
30	268	36.6	355	1	PCT-US01-01239-1145	Sequence 1145, Ap
31	268	36.6	355	21	US-09-764-902-1145	Sequence 1145, Ap
32	255	34.8	309	1	PCT-US01-01332-615	Sequence 615, App
33	255	34.8	309	21	US-09-764-875-615	Sequence 615, App
34	243	33.2	135	26	US-60-196-718-4486	Sequence 4486, Ap
35	238	32.5	208	21	US-09-760-475-2126	Sequence 2126, Ap
36	236	32.2	207	21	US-09-760-443-1321	Sequence 1321, Ap
37	236	32.2	207	21	US-09-760-475-3255	Sequence 3255, Ap
38	236	32.2	207	21	US-09-760-498-920	Sequence 920, App
39	222	30.3	138	1	PCT-US01-01239-1688	Sequence 1688, Ap
40	222	30.3	138	1	PCT-US01-01332-938	Sequence 938, App
41	222	30.3	138	21	US-09-764-875-938	Sequence 938, App
42	222	30.3	138	21	US-09-764-902-1688	Sequence 1688, Ap
43	219.5	29.1	175	1	PCT-US01-03800A-1802	Sequence 1802, Ap
44	199	27.9	107	26	US-60-196-718-5358	Sequence 5358, Ap
45	198	27.0	106	26	US-60-196-710-5336	Sequence 5336, Ap

ALIGNMENTS

RESULT 1
US-09-119-209-2
; Sequence 2, Application US/09119209
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNICK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/119,209
FILING DATE: 20-Jul-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/513278
FILING DATE: 10-AUG-1995

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-119-209-2
```

```

Query Match          100.0%; Score 733; DB 15; Length 372;
Best Local Similarity 100.0%; Pred. No. 6.6e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 197 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 256
QY 61 IQCEPLSAPDLGINMCHPLASFSTSACTFICSEGTLLIGKKTICSSGIMWNSPSPIC 120
    |||||||
DB 257 IQCEPLSAPDLGINMCHPLASFSTSACTFICSEGTLLIGKKTICSSGIMWNSPSPIC 316
QY 121 OKLDKSFMSIKE 132
    |||||||
DB 317 OKLDKSFMSIKE 328
```

```

RESULT 2
US-09-760-443-1437
; Sequence 1437, Application US/09760443
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ12
; CURRENT APPLICATION NUMBER: US/09/760,443
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2164
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1437
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-443-1437
```

```

Query Match          99.2%; Score 727; DB 21; Length 184;
Best Local Similarity 99.2%; Pred. No. 1.2e-62;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 9 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 68
QY 61 IQCEPLSAPDLGINMCHPLASFSTSACTFICSEGTLLIGKKTICSSGIMWNSPSPIC 120
    |||||||
DB 69 IQCEPLSAPDLGINMCHPLASFSTSACTFICSEGTLLIGKKTICSSGIMWNSPSPIC 128
QY 121 OKLDKSFMSIKE 132
```

```

DB 129 OKLDKSFMSIKE 140
|||||
```

```

RESULT 3
US-09-760-475-3252
; Sequence 3252, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ49
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3252
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-3252
```

```

Query Match          99.2%; Score 727; DB 21; Length 184;
Best Local Similarity 99.2%; Pred. No. 1.2e-62;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 9 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 68
QY 61 IQCEPLSAPDLGINMCHPLASFSTSACTFICSEGTLLIGKKTICSSGIMWNSPSPIC 120
    |||||||
DB 69 IQCEPLSAPDLGINMCHPLASFSTSACTFICSEGTLLIGKKTICSSGIMWNSPSPIC 128
QY 121 OKLDKSFMSIKE 132
    |||||||
DB 129 OKLDKSFMSIKE 140
```

```

RESULT 4
PCT-US01-26675-3
; Sequence 3, Application PC/TUS0126675
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Anastasio, Alison E
; APPLICANT: Bieglecki, Karyn M
; APPLICANT: Kilem, Stefanie E
; APPLICANT: Koshiy, Beena
; APPLICANT: Kumar, Anant Madan
; TITLE OF INVENTION: HAPLOTYPES OF THE SELL GENE
; FILE REFERENCE: SELL MWH116-PCT
; CURRENT APPLICATION NUMBER: PCT/US01/26675
; CURRENT FILING DATE: 2001-08-27
; Prior application data removed - refer to PALM or file wrapper
; PRIOR APPLICATION NUMBER: 60/228,262
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 3
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-26675-3
```

```

Query Match          99.2%; Score 727; DB 1; Length 372;
Best Local Similarity 99.2%; Pred. No. 2.5e-62;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
    |||||||
DB 197 CEPLAPELLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 256
```

Qy 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 120
|||||
Db 257 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 316
Qy 121 OKLDKSF5MIKE 132
|||||
Db 317 OKLDKSF5MIKE 328

RESULT 5

US-09-760-475-2123
; Sequence 2123, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2123
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-2123

Query Match 99.2%; Score 727; DB 21; Length 382;
Best Local Similarity 99.2%; Pred. No. 2,1e-62;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 60
|||||
Db 207 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 266
Qy 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 120
|||||
Db 267 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 326
Qy 121 OKLDKSF5MIKE 132
|||||
Db 327 OKLDKSF5MIKE 338

RESULT 6

US-60-207-315-428
; Sequence 428, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-428

Query Match 99.2%; Score 727; DB 26; Length 1078;
Best Local Similarity 99.2%; Pred. No. 8,1e-62;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 60
|||||
Db 779 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 838

Qy 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 120
|||||
Db 839 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 898
Qy 121 OKLDKSF5MIKE 132
|||||
Db 899 OKLDKSF5MIKE 910

RESULT 7

US-60-212-659-523
; Sequence 523, Application US/60212659
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1000674
; CURRENT APPLICATION NUMBER: US/60/212,659
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 879
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-212-659-523

Query Match 99.2%; Score 727; DB 26; Length 1078;
Best Local Similarity 99.2%; Pred. No. 8,1e-62;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 60
|||||
Db 779 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 838
Qy 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 120
|||||
Db 839 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 898
Qy 121 OKLDKSF5MIKE 132
|||||
Db 899 OKLDKSF5MIKE 910

RESULT 8

US-60-230-435-1751
; Sequence 1751, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1751
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1751

Query Match 99.2%; Score 727; DB 26; Length 1078;
Best Local Similarity 99.2%; Pred. No. 8,1e-62;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 60
|||||
Db 779 CEPLAEPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 838

QY 61 IOCEPLASAPDLGIMNCSHPLASFSTACTICSECTELIGKKTICSSGIMNSPSPIC 120
Db 839 IOCEPLASAPDLGIMNCSHPLASFSTACTICSECTELIGKKTICSSGIMNSPSPIC 898
QY 121 QKLDKSFMSIKE 132
Db 899 QKLDKSFMSIKE 910

RESULT 9

PCT-US92-03970-2
Sequence 2, Application PC/TUS9203970
GENERAL INFORMATION:
APPLICANT: Dana-Farber Cancer Institute, Inc.
TITLE OF INVENTION: LEUKOCYTE-ASSOCIATED CELL SURFACE
PROTEIN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03970
FILING DATE: 19920513
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Helne, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DCT-152Bq9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-03970-2

Query Match 97.3%; Score 713; DB 1; Length 385;
Best Local Similarity 97.0%; Pred. No. 6.1e-61;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAEPLGMDCTHFGNFSFSSQCAFSCSEGTNLGTIEETTCGPFQWMSSEPTCOV 60
Db 210 CEPLAEPLGMDCTHFGNFSFSSQCAFSCSEGTNLGTIEETTCGPFQWMSSEPTCOV 269
QY 61 IOCEPLASAPDLGIMNCSHPLASFSTACTICSECTELIGKKTICSSGIMNSPSPIC 120
Db 270 IOCEPLASAPDLGIMNCSHPLASFSTACTICSECTELIGKKTICSSGIMNSPSPIC 329
QY 121 QKLDKSFMSIKE 132
Db 330 QKLDKSFMSIKE 341

RESULT 10
PCT-US94-00909-2
Sequence 2, Application PC/TUS9400909
GENERAL INFORMATION:
APPLICANT:

TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS BLOCKING
AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00909
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/962,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/737,092
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-00909-2

Query Match 97.3%; Score 713; DB 1; Length 385;
Best Local Similarity 97.0%; Pred. No. 6.1e-61;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAEPLGMDCTHFGNFSFSSQCAFSCSEGTNLGTIEETTCGPFQWMSSEPTCOV 60
Db 210 CEPLAEPLGMDCTHFGNFSFSSQCAFSCSEGTNLGTIEETTCGPFQWMSSEPTCOV 269
QY 61 IOCEPLASAPDLGIMNCSHPLASFSTACTICSECTELIGKKTICSSGIMNSPSPIC 120
Db 270 IOCEPLASAPDLGIMNCSHPLASFSTACTICSECTELIGKKTICSSGIMNSPSPIC 329
QY 121 QKLDKSFMSIKE 132
Db 330 QKLDKSFMSIKE 341

RESULT 11
US-08-008-459-2
Sequence 2, Application US/08008459
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston

```
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/008,459
FILING DATE: 25-JAN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/862,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/737,092
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-318XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-008-459-2

Query Match          97.3%: Score 713; DB 4; Length 385;
Best Local Similarity 97.0%: Pred. No. 6,1e-61;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/862,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-318XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539-2

Query Match          97.3%: Score 713; DB 7; Length 385;
Best Local Similarity 97.0%: Pred. No. 6,1e-61;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

Db 330 QKLDKSFMSIKE 341

RESULT 13

US-08-410-569-2

Sequence 2 Application US/08410569

GENERAL INFORMATION:

APPLICANT: Tedder, Thomas F.

APPLICANT: Spertini, Olivier G.

TITLE OF INVENTION: LEUKOCYTE ADHESION MOLECULE-1 (LAM-1)

TITLE OF INVENTION: AND LIGAND THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes

STREET: Ten Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/410,569

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/770,608

FILING DATE: 03-OCT-1991

APPLICATION NUMBER: US 07/700,773

FILING DATE: 15-MAY-1991

ATTORNEY/AGENT INFORMATION:

NAME: Heine, Holliday C.

REGISTRATION NUMBER: 34,346

REFERENCE/DOCKET NUMBER: DFCG-152EX

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-2290

TELEFAX: (617) 451-0313

TELEX: 940675

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 385 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-410-569-2

Query Match 97.3%; Score 713; DB 8; Length 385;
Best Local Similarity 97.0%; Pred. No. 6.1e-61;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPRGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 210 CEPLAPELGTMDCTHPRGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 269
QY 61 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTKTICSSGIMNSPSPIC 120
Db 270 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTKTICSSGIMNSPSPIC 329
QY 121 QKLDKSFMSIKE 132
Db 330 QKLDKSFMSIKE 341

RESULT 14

US-09-758-449-1158

Sequence 1158, Application US/09758449
GENERAL INFORMATION:
APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PM026
CURRENT APPLICATION NUMBER: US/09/758,449
CURRENT FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
NUMBER OF SEQ ID NOS: 1478
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1158
LENGTH: 341
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (215)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-449-1158

Query Match 92.0%; Score 674; DB 21; Length 341;
Best Local Similarity 99.2%; Pred. No. 3.5e-57;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPRGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 219 CEPLAPELGTMDCTHPRGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 278
QY 61 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTKTICSSGIMNSPSPIC 120
Db 279 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTKTICSSGIMNSPSPIC 338
QY 121 Q 121
Db 339 Q 339

RESULT 15
US-09-760-443-1328
Sequence 1328, Application US/09760443
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PZ12
CURRENT APPLICATION NUMBER: US/09/760,443
CURRENT FILING DATE: 2001-01-16
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2164
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1328
LENGTH: 341
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (215)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-443-1328

Query Match 92.0%; Score 674; DB 21; Length 341;
Best Local Similarity 99.2%; Pred. No. 3.5e-57;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPRGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
Db 219 CEPLAPELGTMDCTHPRGNFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 278
QY 61 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTKTICSSGIMNSPSPIC 120
Db 279 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTKTICSSGIMNSPSPIC 338

OY 121 Q 121
DB 339 Q 339

```
RESULT 16
US-09-119-209-4
; Sequence 4, Application US/09119209
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,209
; FILING DATE: 20-Jul-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513278
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-119-209-4

Query Match          70.7%; Score 518; DB 15; Length 372;
Best Local Similarity 68.2%; Pred. No. 6, 6e-42;
Matches 90; Conservative 13; Mismatches 29; Indels 0; Gaps 0;
```

DB 317 QETNRSFSKIKE 328

```
RESULT 17
PCT-US94-09395-4
; Sequence 4, Application PC/TUS9409395
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: Expression Control Sequences of the P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09395
; FILING DATE: 19-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 830 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US94-09395-4

Query Match          43.0%; Score 315; DB 1; Length 830;
Best Local Similarity 43.9%; Pred. No. 1, 1e-21;
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;
```

```

1      STATE: GA
2      COUNTRY: USA
3      ZIP: 30309-3450
4
5      COMPUTER READABLE FORM:
6      MEDIUM TYPE: floppy disk
7      COMPUTER: IBM PC compatible
8      OPERATING SYSTEM: PC-DOS/MS-DOS
9      SOFTWARE: PatentIn Release #1.0, Version #1.25
10     CURRENT APPLICATION DATA:
11     APPLICATION NUMBER: US/08/449,687B
12     FILING DATE: 24-MAY-1995
13     CLASSIFICATION: 435
14
15     PRIOR APPLICATION DATA:
16     APPLICATION NUMBER: US 08/110,158
17     FILING DATE: 20-AUG-1993
18
19     PRIOR APPLICATION DATA:
20     APPLICATION NUMBER: US 07/320,408
21     FILING DATE: 08-MAR-1989
22     ATTORNEY/AGENT INFORMATION:
23     NAME: Pabst, Patricia L.
24
25     REGISTRATION NUMBER: 31,284
26     TELECOMMUNICATION INFORMATION:
27     TELEPHONE: (404)873-8794
28     TELEFAX: (404)873-8795
29
30     INFORMATION FOR SEQ ID NO: 4:
31     SEQUENCE CHARACTERISTICS:
32     LENGTH: 830 amino acids
33     TYPE: amino acid
34     STRANDEDNESS: single
35     TOPOLOGY: linear
36
37     MOLECULE TYPE: peptide
38     US-08-449-687B-4

```

Query Match	43.0%;	Score 315;	DB 8;	Length 830;
Best Local Similarity	43.9%;	Pred. No. 1.1e-21;		
Matches 54;	Conservative 21;	Mismatches 48;	Indels 0;	Gaps 0;

QY 1 CEPLEAPLGTMDCTHPGNFSFSSCOAFSCSECTNLIGLEETTCGFGNMSSEPPICOV 60
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 324 CQHLEAFSEGTMDCVHFLTLTAAYVSSCKFEQPGYRVRGIDMLRCDISGHRSAFLPICEA 383
QY 61 IQCEPLAPDLGINNCSPILPLASFSFTACFTICSEBTELLIGKKTITESSGIMSNSPPTIC 120
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 384 ITCSEPLPVHGSMDCSPLTAFOYDTNCSFRCAEGMLRGAADVRCDDLNGQWTAAPVC 443
QY 121 QKL 123
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
444 QAL 446
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db

```

RESULT 19
US-10-020-141-10
; Sequence 10, Application US/10020141.
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Ableson, Allen
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MM-002
; CURRENT APPLICATION NUMBER: US/10/020,141
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/313,097
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/327,485
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 830
; TYPE: PRF
; ORGANISM: Homo sapiens
; US-10-020-141-10

```

Query Match	43.0%;	Score 315;	DB 24;	Length 830;
Best Local Similarity	43.9%;	Pred. No. 1.1e-21;		
Matches 54;	Conservative 21;	Mismatches 48;	Indels 0;	Gaps 0;

```

0Y      1  CEPLEAPBELGMDCHHPGNGNFSFSQQAFCFSCSEGNLTGIEFTTCGFGNGNNSPFPFCOY  60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      324  COHLEAPBEGMDCHPLTAAAYSSCCFECQDPQGRVANGDMLRCIDSGHMSAPLPTCEA  383
0Y      61  IQCEPLSAPDGIWNMCSHPLASFPSYACFTGSEGTELLGKKTTIOESSGIMWSPDIC  120
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      384  ISCEPLESPVHGSMDCSPSLRAFOFDYTCSEFCAEGMLRGADIVRCIDNLGOWTAPAPVC  443
0Y      121  OKL  123
        |||
Db      444  QAL  446

```

```

RESULT 20
US-60-207-315-467
; Sequence 467, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 467
; LENGTH: 700
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(700)
; OTHER INFORMATION: xaa = Any Amino Acid
US-60-207-315-467

```

Query Match	41.7%;	Score 306;	DB 26;	Length 700;
Best Local Similarity	41.3%;	Pred. No. 6.7e-21;		
Matches 50;	Conservative 23;	Mismatches 48;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 21
PCT-US99-28965-19
: Sequence 19, Application PC/TUS9928965
: GENERAL INFORMATION:
: APPLICANT: Monla, Brett P.
: APPLICANT: Xu, Xiaoxing S.
: APPLICANT: Isis Pharmaceuticals, Inc.
: TITLE OF INVENTION: METHODS OF MOULATING TUMOR NECROSIS FACTOR
: TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
: FILE REFERENCE: ISPH-0424
: CURRENT APPLICATION NUMBER: PCT/US99/28965
: CURRENT FILING DATE: 1999-12-08
: EARLIER APPLICATION NUMBER: US 09/209,668
: EARLIER FILING DATE: 1998-12-10

```


OY 121 Q 121
Db 300 K 300

RESULT 24
US-09-266-091-2
Sequence 2, Application US/09266091
GENERAL INFORMATION:
APPLICANT: Klimuk, Sandra K
APPLICANT: Semple, Sean C
APPLICANT: Scherrier, Peter
APPLICANT: Hope, Michael J.
APPLICANT: Bennett, C. Frank
TITLE OF INVENTION: ENHANCED EFFICACY OF LIPOSOMAL ANTISENSE THERAPY
FILE REFERENCE: ISPH-0342
CURRENT APPLICATION NUMBER: US/09/266,091
CURRENT FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 08/657,753
PRIOR FILING DATE: 1996-05-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 610
TYPE: PRT
ORGANISM: Homo sapiens
US-09-266-091-2

Query Match
Best Local Similarity 40.4%; Score 296; DB 16; Length 610;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

OY 1
Db 180
OY 61
Db 240
OY 121 Q 121
Db 300 K 300

RESULT 25
US-09-266-091A-2
Sequence 2, Application US/09266091A
GENERAL INFORMATION:
APPLICANT: Klimuk, Sandra K.
APPLICANT: Semple, Sean C.
APPLICANT: Scherrier, Peter
APPLICANT: Hope, Michael J.
TITLE OF INVENTION: Enhanced Efficacy of Liposomal Antisense
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,091A
FILING DATE: 10-Mar-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/657,753
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kezer, William B.
REGISTRATION NUMBER: 37,369
REFERENCE/DOCKET NUMBER: 16303-003600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-266-091A-2

Query Match
Best Local Similarity 40.4%; Score 296; DB 16; Length 610;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

OY 1
Db 180
OY 61
Db 240
OY 121 Q 121
Db 300 K 300

RESULT 26
US-09-784-356-122
Sequence 122, Application US/09784356
GENERAL INFORMATION:
APPLICANT: Murray, Richard
APPLICANT: Glynn, Richard
APPLICANT: Watson, Susan R.
APPLICANT: EOS Biotechnology, Inc.
TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,
TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
FILE REFERENCE: 018501-00071005
CURRENT APPLICATION NUMBER: US/09/784,356
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: US 60/148,425
PRIOR FILING DATE: 1999-08-11
PRIOR APPLICATION NUMBER: US 09/637,977
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 135
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 122
LENGTH: 610
TYPE: PRT
ORGANISM: Homo sapiens
US-09-784-356-122

Query Match
Best Local Similarity 40.4%; Score 296; DB 21; Length 610;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

OY 1
Db 180
OY 61
Db 240
OY 121 Q 121
Db 300 K 300


```

; NAME/KEY: SITE
; LOCATION: (113)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (149)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (150)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (151)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (193)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
PCT-US01-01239-1145
```

```
Query Match
Best Local Similarity 36.6%; Score 268; DB 1; Length 355;
Matches 48; Conservative 17; Mismatches 56; Indels 0; Gaps 0;
```

```
OY 1 CEPLEAPELGTMDCTHPGNGFSFSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60
    | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 35 CTPILSPONGTWTQVQPLGSSSYSTQCFICDEGYLSGPERLDCTRSGRWTDSPMCEA 94
OY 61 IQCEPLSAPDLGIMNCSPHLASFSTSACTFICSEGTLLIGKRTKTCSSGIMNSPSPIC 120
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 95 IKCPFLFAPEQGSIDCSXRGEPFNVGSTCHFSNNNGFLEBPNNVECTSGRWSXXXPTC 154
OY 121 Q 121
Db 155 K 155
```

```
RESULT 31
US-09-764-902-1145
; Sequence 1145, Application US/09764902
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213
; CURRENT APPLICATION NUMBER: US/09/764,902
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1145
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (113)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (149)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (150)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (151)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (193)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-764-902-1145
```

```
Query Match
Best Local Similarity 36.6%; Score 268; DB 21; Length 355;
Matches 48; Conservative 17; Mismatches 56; Indels 0; Gaps 0;
```

```
OY 1 CEPLEAPELGTMDCTHPGNGFSFSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60
    | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 35 CTPILSPONGTWTQVQPLGSSSYSTQCFICDEGYLSGPERLDCTRSGRWTDSPMCEA 94
OY 61 IQCEPLSAPDLGIMNCSPHLASFSTSACTFICSEGTLLIGKRTKTCSSGIMNSPSPIC 120
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 95 IKCPFLFAPEQGSIDCSXRGEPFNVGSTCHFSNNNGFLEBPNNVECTSGRWSXXXPTC 154
OY 121 Q 121
Db 155 K 155
```

```
RESULT 32
PCT-US01-01332-615
; Sequence 615, Application PC/TUS0101332
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01332
; CURRENT FILING DATE: 2001-05-09
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 615
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01332-615
```

```
Query Match
Best Local Similarity 34.8%; Score 255; DB 1; Length 309;
Matches 47; Conservative 16; Mismatches 58; Indels 0; Gaps 0;
```

```
OY 1 CEPLEAPELGTMDCTHPGNGFSFSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60
    | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 121 CPALTTCGGTMCRRHHPGTFGRNTTCFPCNMGFTLLIGSTISCRSPGQWAVTACRA 180
OY 61 IQCEPLSAPDLGIMNCSPHLASFSTSACTFICSEGTLLIGKRTKTCSSGIMNSPSPIC 120
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 181 VKSELHVNRPIAMNCNLMGNFSYGSICFHLLEGOLLNGSAQTACQENGHMSTVTPIC 240
OY 121 Q 121
Db 241 Q 241
```

```
RESULT 33
US-09-764-875-615
; Sequence 615, Application US/09764875
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 615
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-615
```

```
Query Match
Best Local Similarity 34.8%; Score 255; DB 21; Length 309;
Matches 47; Conservative 16; Mismatches 58; Indels 0; Gaps 0;
```



```

/ SOFTWARE:patentln Ver. 2.0
/ SEQ ID NO 920
/ LENGTH: 207
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (151)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (185)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (191)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (194)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (198)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-498-920

Query Match          32.2%; Score 236; DB 21; Length 207;
Best Local Similarity 37.4%; Pred. No. 1.2e-14;
Matches 46; Conservative 17; Mismatches 54; Indels 6; Gaps 3

OY      1 CEPLAEPLGMOCTH-PRGNFSPSSOCAPSCSEGNLTGIEFTTCGFGMMSSPEPTCQ 59
DB      23 CDAAHOPPKGLVRCAMHSPIGEFTYKSSCAPSCSEGPLHGSTOLCTSGOGWTEERVPCQ 82

OY      60 VIOCEPLSAPDLGIMNCS-HPLASFSFTSACTFTICSEGTETLIGKKRTICESGSIWNSP 118
DB      83 VVKCSSLAVPGKIMSSGSEPV---FGTVCKFACEGWTLNGSARPCGATGHHSGILP 138

OY      119 ICQ 121
DB      139 TCE 141

RESULT 39
PCT-US01-01239-1688
/ Sequence 1688, Application PC/TUS0101239
/ GENERAL INFORMATION:
/ APPLICANT: Human Genome Sciences, Inc., et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT213PCT
/ CURRENT APPLICATION NUMBER: PCT/US01/01239
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - refer to PAM or file wrapper
/ NUMBER OF SEQ ID NOS: 2318
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 1688
/ LENGTH: 138
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (31)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (76)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (127)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (135)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01239-1688

```


This Page blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:43 : Search time 28.63 Seconds
(without alignments)
112.615 Million cell updates/sec

Title: US-09-119-209-2_COPY_197_328

Perfect score: 733
Sequence: 1 CEPLEAPELGTMDCTHPRGN.....WSNPSPICKKIDKSRSMIKE 132

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	733	100.0	372	2 US-08-513-278-2	Sequence 2, Appl1
2	733	100.0	372	6 5514582-2	Patent No. 5514582
3	713	97.3	385	1 US-08-340-539A-2	Sequence 2, Appl1
4	713	97.3	385	2 US-08-461-592B-2	Sequence 2, Appl1
5	518	70.7	372	2 US-08-513-278-4	Sequence 4, Appl1
6	518	70.7	372	6 5514582-4	Patent No. 5514582
7	487	66.4	126	6 5514582-31	Patent No. 5514582
8	315	43.0	830	1 US-08-110-158-4	Sequence 4, Appl1
9	314	42.8	830	5 PCT-US91-05059-2	Sequence 2, Appl1
10	305	41.4	574	6 5378464-3	Patent No. 5378464
11	303.5	41.4	830	6 5378464-2	Patent No. 5378464
12	296	40.4	610	1 US-08-365-470-3	Sequence 3, Appl1
13	296	40.4	610	3 US-09-209-668-19	Sequence 19, Appl1
14	296	40.4	610	4 US-09-009-490A-89	Sequence 89, Appl1
15	296	40.4	610	6 5217870-2	Patent No. 5217870
16	245	33.4	484	2 US-08-252-193C-9	Sequence 9, Appl1
17	245	33.4	484	3 US-09-276-197-9	Sequence 9, Appl1
18	165.5	22.6	240	3 US-08-824-692-23	Sequence 23, Appl1
19	163.5	22.3	127	6 5514582-32	Patent No. 5514582
20	163.5	22.3	1466	6 5256642-6	Patent No. 5256642
21	163.5	22.3	1466	6 5472939-6	Patent No. 5472939
22	163.5	22.3	1537	6 5256642-5	Patent No. 5256642
23	163.5	22.3	1537	6 5472939-5	Patent No. 5472939
24	163.5	22.3	1847	6 5256642-10	Patent No. 5256642
25	163.5	22.3	1847	6 5472939-10	Patent No. 5472939
26	163.5	22.3	2039	6 5256642-2	Patent No. 5256642
27	163.5	22.3	2039	6 5472939-2	Patent No. 5472939

ALIGNMENTS

28	162	22.1	128	6 5514582-33	Patent No. 5514582
29	153	20.9	62	1 US-08-202-047-20	Sequence 20, Appl1
30	153	20.9	62	3 US-08-964-690-20	Sequence 2, Appl1
31	152.5	20.8	363	4 US-08-961-234B-2	Sequence 2, Appl1
32	149	20.3	216	6 US-08-824-692-24	Sequence 24, Appl1
33	148.5	20.3	126	6 5514582-35	Patent No. 5514582
34	140.5	19.2	265	2 US-08-177-109A-57	Sequence 57, Appl1
35	140.5	19.2	265	2 US-08-687-706-57	Sequence 57, Appl1
36	140.5	19.2	764	2 US-08-177-109A-2	Sequence 2, Appl1
37	140.5	19.2	764	2 US-08-687-706-2	Sequence 2, Appl1
38	139	19.0	126	6 5514582-43	Patent No. 5514582
39	139	19.0	263	1 US-07-906-983-2	Sequence 2, Appl1
40	139	19.0	323	2 US-08-435-149-2	Sequence 2, Appl1
41	139	19.0	324	1 US-08-310-416A-14	Sequence 14, Appl1
42	139	19.0	324	1 US-08-888-171-14	Sequence 14, Appl1
43	139	19.0	577	2 US-08-435-149-3	Sequence 3, Appl1
44	139	19.0	611	4 US-09-475-460A-32	Sequence 32, Appl1
45	138	18.8	197	2 US-08-356-361-27	Sequence 27, Appl1

RESULT 1
Sequence 2, Application US/08513278
Patent No. 5840844
GENERAL INFORMATION:
APPLICANT: LASKY, LAURENCE A.
APPLICANT: STACHELL, SCOTT E.
APPLICANT: ROSEN, STEVEN D.
APPLICANT: SINGER, MARK S.
APPLICANT: YEDNOCK, TED A.
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION NUMBER: US/08/513,278
FILING DATE: 10-AUG-1995
CLASSIFICATION: 5530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/059027
FILING DATE: 06-MAY-1993
APPLICATION NUMBER: 07/786149
FILING DATE: 31-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/315015
FILING DATE: 23-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 56501C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-513-278-2

Query Match	100.0%;	Score 733;	DB 2;	length 372;
Best Local Similarity	100.0%;	Pred. NO. 4.1e-71;		
Matches 132;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY 1 CEPLAPELLGMDCTHPGFNFSFSSQCAFSCSEGNLTGLETTCPFGNWSSPEPTQY 60
 |||||
Db 197 CEPLEAPELGTMDCTHPGFNFSSQCAFSCSEGNLTGLETTCPFGNWSSPEPTQY 256

QY 61 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTELIGKKTICCESSIWSNPSIC 120
 Db 257 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTELIGKKTICCESSIWSNPSIC 316

```
QY      121 QKLDKSFMSIKE 132
        |||||
Db      317 QKLDKSFMSIKE 328
```

RESULT 2
5514582-2
; Patent No. 5514582

```

: TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
: IMMUNOGLOBULINS
: NUMBER OF SEQUENCES: 43

```

APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:

```

;SEQ ID NO:2
;      LENGTH: 372
5514582-2

```

Query Match	100.0%;	Score 733;	DB 6;	length 372;
Best Local Similarity	100.0%;	Pred. No. 4.1e-71;		
Matches 132;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

QY 61 ICEPLSADDLGINMCSHPLASFSTACTFICSEGTELIGKKKTCICSSGIMSNPSPIC 120C
|||||
Db 257 IQCEPLSADDLGINMCSHPLASFSTACTFICSEGTELIGKKKTCICSSGIMSNPSPIC 316C

QY	121	QKDKSFSMIKE	132
Db	317	QKDKSFSMIKE	328

RESULT 3
US-08-340-539A-2

Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder
ADDRESSEE: Z

TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas

```

; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:

```

```
;
;      MEDIUM TYPE: Floppy disk
;
;      COMPUTER: IBM PC compatible
;
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;
;      Copyright © 1987 PatentIn
```

; APPLICATION NUMBER: US/08/340,539A
 ; FILING DATE: 16-NOV-1994
 ; CLASSIFICATION: 514
 ; PRIORITY ADDITION: DATA

; APPLICATION NUMBER: US 08/008,459
 ; FILING DATE: 25-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Curren, Todd

REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
ATTORNEY: 313. 506. 0000

TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

```

;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
nc-09-240-520-3

```

Query Match	97.3%	Score 713	DB 1	Length 385
Best Local Similarity	97.0%	Pred. No. 6e-69		
Matches 128	Conservative 2	Mismatches 2	Indels 0	Gaps 0

QY 1 CEELEAPELGTMDCTHPFGNFSFSSQCAFSCSEETNLGTGIEETTCGPFGNSSSPEPICQV 60
|||||
|||:
210 CEELEAPELGTMDCTHPFGNFSFSSQCAFSCSEETNLGTGIEETTCGPFGNSSSPEPICQV 269

QY 61 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTelikKKKTCICSSSGIWSNPSPIC 120
Db 270 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTelikKKKTCICSSSGIWSNPSPIC 329

```
QY      121 QKDKSFSMIKE 132
          |||||
Db      330 QKDKSFSMIKE 341
```

RESULT 4
US-08-461-592B-2

Patent No. 5834425
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
ADDRESSEE: Kenneth C. Coffey, S.

; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
 ; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
 ;
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:

```

;
;      SOFTWARE: Patent Release #1.0, V
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/461,592B

```

```

: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/340,539
: FILING DATE: 16-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/008,459
: FILING DATE: 25-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: James F. Haley, Jr.
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: CG-104
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 596-9000
: TELEFAX: (212) 596-9090
: TELEX: 14-8367
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 385 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-461-592B-2

```

Query Match	97.38;	Score 713;	DB 2;	Length 385;
Best Local Similarity	97.08;	Pred. No. 6e-69;		
Matches 128;	Conservative	2;	Mismatches 2;	Indels 0;
			Gaps	0;

Qy	1	CEPLEAPRLGTMDCDTHPRGNGSFSFQCAFSCSEGTNLGIEETTCGFGNMSSPEPTQY	60
Db	210	CEPLEAPRLGTMDCDTHPLGNFNFRSQCAFSCSEGTNLGIEETTCGFGNMSSPEPTQY	269
Qy	61	IQCEPLSPDGLGNNCCHPLSPSFTSACPTICSEGNELIQKKTICDESSGIMNSPPTC	120
Db	270	IQCEPLSPDGLGNNCCHPLSPSFTSACPTICSEGNELIQKKTICDESSGIMNSPPTC	329
Qy	121	QKLDKSFMSIKE	132
Db	330	QKLDKSFMSIKE	341

US-RESULT 5
US-08-513-278-4
Sequence 4, Application US/08513278
Patent No. 5840844
GENERAL INFORMATION:
APPLICANT: LASKY, LAURENCE A.
APPLICANT: STACHELL, SCOTT E.
APPLICANT: ROSEN, STEVEN D.
APPLICANT: SINGER, MARK S.
APPLICANT: YEDNOCK, TED A.
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 Inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,278
FILING DATE: 10-AUG-1995
CLASSIFICATION: 5530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/059027
FILING DATE: 06-MAY-1993

APPLICATION NUMBER: 07/786149
FILING DATE: 33-OCY-1991
PRIOR APPLICATION DATA: 07/315015
APPLICATION NUMBER: 07/315015
FILING DATE: 23-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 33,055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-513-278-4

Query Match	70.7%;	Score 518;	DB 2;	Length 372;
Best Local Similarity	68.2%;	Pred. No. 5.4e-48;		
Matches	90;	Conservative 13;	Mismatches 29;	Indels 0;
			Gaps	0;

```
Qy 1 CEELEAPBLGTMDCTHPPGNGNSFSSOCACFSCSEBNTNLTGIEHTTGCFFGNMSSPEPICOY 60
Db 197 CEELEAPBLGTMDCIHPHGNFSFOSKCAFNCSEBRELTLGTAETOCGAGNMSSPEPICOY 256
Qy 61 IQCEPLSAPDLGINNCMSHPLASFSPFTSACPTICSEBTELLGKKTTIESSGIMWNSPPIIC 120
Db 257 VQCELEAPBLGTMDCIHPHGNFSFOSKCAFNCEBRELTLGTAETOCGAGNMSSPEPIC 316
Qy 121 OKLDKSFEMIKE 132
Db 317 QETNNSFEMIKE 328
```

RESULT - 6
 5514582-4
 Patent No. 5514582
 APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
 TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
 IMMUNOGLOBULINS
 NUMBER OF SEQUENCES: 43
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/185,670
 FILING DATE: 21-JAN-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 986,931
 FILING DATE: 08-DEC-1992
 APPLICATION NUMBER: 808,122
 FILING DATE: 16-DEC-1991
 APPLICATION NUMBER: 440,625
 FILING DATE: 22-NOV-1989
 APPLICATION NUMBER: 315,015
 FILING DATE: 23-FEB-1989
 SEQ ID NO:4:
 LENGTH: 372
 5514582-4

Query Match	70.7%	Score 518;	DB 6;	Length 372;
Best Local Similarity	68.2%	Pred. No. 5.4e+48;		
Matches	90;	Conservative 13;	Mismatches 22;	Indels 0; Gaps 0;
QY	1	CEPLEAPELGMDCHHPGNGNFSFSOCAFSCSEGTNLGIEETTGCPGNWSSPEPTQY	60	
Db	197	CEPLEAPELGMDCHHPGNGNFSFSOCAFSCSEGTNLGIEETTGCPGNWSSPEPTQY	256	
QY	61	IQCEPLAPADLGIMNCNCHPLASFSTACTCTCGSEGTLLGKRRKTCIBSSSGIWNSPIC	120	
Db	257	VQCEPLAPADLGIMNCNCHPLASFSTACTCTCGSEGTLLGKRRKTCIBSSSGIWNSPIC	316	

TELEFAX: 404-572-6555	NAME/KEY: Disulfide-bond
INFORMATION FOR SEQ ID NO: 2:	LOCATION: 461..474
SEQUENCE CHARACTERISTICS:	FEATURE:
LENGTH: 830 amino acids	NAME/KEY: Disulfide-bond
TYPE: AMINO ACID	LOCATION: 478..505
STRANDEDNESS: single	FEATURE:
TOPOLOGY: linear	NAME/KEY: Disulfide-bond
MOLECULE TYPE: peptide	LOCATION: 510..554
HYPOTHETICAL: YES	FEATURE:
ANTI-SENSE: YES	NAME/KEY: Disulfide-bond
FRAGMENT TYPE: N-terminal	LOCATION: 523..536
ORIGINAL SOURCE:	FEATURE:
ORGANISM: Homo sapien	NAME/KEY: Disulfide-bond
TISSUE TYPE: Blood	LOCATION: 540..567
CELL TYPE: Endothelial	FEATURE:
FEATURE:	NAME/KEY: Disulfide-bond
NAME/KEY: Disulfide-bond	LOCATION: 572..616
LOCATION: 4..25	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 60..158	LOCATION: 585..616
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 131..150	LOCATION: 602..629
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 163..174	LOCATION: 642..686
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 168..183	LOCATION: 655..668
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 185..194	LOCATION: 672..699
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 200..244	LOCATION: 704..748
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 213..226	LOCATION: 717..730
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 230..257	LOCATION: 734..761
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Binding-site
LOCATION: 262..306	LOCATION: 54
FEATURE:	OTHER INFORMATION: /note- "Potential asparagine-linked
NAME/KEY: Disulfide-bond	OTHER INFORMATION: glycosylation site"
LOCATION: 275..288	FEATURE:
FEATURE:	NAME/KEY: Binding-site
NAME/KEY: Disulfide-bond	LOCATION: 98
LOCATION: 292..319	OTHER INFORMATION: /note- "Potential asparagine-linked
FEATURE:	OTHER INFORMATION: glycosylation site"
NAME/KEY: Disulfide-bond	FEATURE:
LOCATION: 324..368	NAME/KEY: Binding-site
FEATURE:	LOCATION: 180
NAME/KEY: Disulfide-bond	OTHER INFORMATION: /note- "Potential asparagine-linked
LOCATION: 337..350	OTHER INFORMATION: glycosylation site"
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Binding-site
LOCATION: 354..381	LOCATION: 212
FEATURE:	OTHER INFORMATION: /note- "Potential asparagine-linked
NAME/KEY: Disulfide-bond	OTHER INFORMATION: glycosylation site"
LOCATION: 386..430	FEATURE:
FEATURE:	NAME/KEY: Binding-site
NAME/KEY: Disulfide-bond	LOCATION: 219
LOCATION: 399..412	OTHER INFORMATION: /note- "Potential asparagine-linked
FEATURE:	OTHER INFORMATION: glycosylation site"
NAME/KEY: Disulfide-bond	FEATURE:
LOCATION: 416..443	NAME/KEY: Binding-site
FEATURE:	LOCATION: 411
NAME/KEY: Disulfide-bond	OTHER INFORMATION: /note- "Potential asparagine-linked
LOCATION: 448..492	OTHER INFORMATION: glycosylation site"
FEATURE:	FEATURE:
	NAME/KEY: Binding-site

US-08-365-470-3


```

Query Match 40.4%; Score 296; DB 6; Length 610;
Best Local Similarity 41.3%; Pred. No.7.5e-24;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0

QY      1 CEELEAPELGTMDCTHPFNFFSSOCARFSCSEGTMLTGIETTCGPFGNWSSPEPTCOV 60
       | :|::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| 
DB      180 CPALSEPGLSVLCSPHLNGFSYNSSCSISCDRGILPSSMETIQQCNSOEWAPIPCANV 239

QY      61 ICFEPLSADDLGMNCSHFLASFSTACTFLICSETELIGKKITICSSGIWSNPSPIC 120
       || ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| 
DB      240 VECDAVTNPNANGFEVECFONPSFPWNNTCTFDCEEGFILMGAOSLOCTSSGNWDNEKPTC 299

QY      121 Q 121

DB      300 K 300

RESULT   16
US-08-252-493C--9
Sequence 9, Application US/08252493C
Patent No. 5891645
GENERAL INFORMATION:
APPLICANT: Rollins, Scott P.
APPLICANT: Rother, Russell P.
APPLICANT: Evans, Mark J.
APPLICANT: Malis, Louis A.
TITLE OF INVENTION: PORCINE E-SELECTIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSSEE: Seth A. Fidel
STREET: 25 Science Park, Box 15
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 Kb storage
COMPUTER: PC compatible
OPERATING SYSTEM: DOS 6.2
SOFTWARE: Wordperfect 6.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252.493C
FILING DATE: June 1, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: predicted amino acid sequence of
DESCRPTION: Porcine E-selectin
US-08-252-493C-9

Query Match 33.4%; Score 245; DB 2; Length 484;
Best Local Similarity 35.0%; Pred. No.1.7e-18;
Matches 43; Conservative 22; Mismatches 56; Indels 2; Caps 2;

1 CEPLAEAPELGTMDCTHPFGNFSSOCARFSCSGTNLTGLETTCCGFPGMWSSPEPTCQ 59
| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| ::|| 

```

```

Db      240  CDIVGHGQNDGAVGNSHSSIGEFAYKSTCHFTCAAGFGLGQPAQIECTAQAQGWQQAIVCK 299
QY      60  VIOCEPLASPDGLGIMNCSH-PLASFSTSACTFICSEGTELIGKKTKTICSSSGIWSNPSP 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      300  AVCPAASQPKNGIYVFTHTSPTEGFTYKSSCAFCSEGFELRGSAQLACTSOGQMTQDEV 359
QY      119  ICQ 121
      ||
Db      360  SCQ 362

RESULT 17
US-09-276-197-9
: Sequence 9, Application US/09276197
: Patent No. 6040428
: GENERAL INFORMATION:
: APPLICANT: Rollins, Scott
: APPLICANT: Rother, Russell P.
: APPLICANT: Evans, Mark J.
: APPLICANT: Mattis, Louis A.
: TITLE OF INVENTION: PORCINE E-SELECTIN
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seth A. Fidel
: STREET: 25 Science Park, Box 15
: CITY: New Haven
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06511
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 750 Kb storage
: COMPUTER: PC compatible
: OPERATING SYSTEM: DOS 6.2
: SOFTWARE: Wordperfect 6.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/276,197
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/252,493
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Fidel, Seth A.
: REGISTRATION NUMBER: 38,449
: REFERENCE/DOCKET NUMBER: ALX-138
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (203) 776-1790
: TELEFAX: (203) 772-3655
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 484 amino acids
: TYPE: amino acid
: STRANDEDNESS: Single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: DESCRIPTION: predicted amino acid sequence of
: DESCRIPTION: Porcine E-selectin
: US-09-276-197-9

Query Match 33.4%; Score 245; DB 3; Length 484;
Best Local Similarity 35.0%; Pred. No. 1,7e-18;
Matches 43; Conservative 22; Mismatches 56; Indels 2; Gaps 2;

QY      1  CEPLEAPELGTMDCH-PFGNFSFSSOCAFCSEGTNLGTGIBETTCGPGNMSPEPICQ 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      240  CDIVGHGQNDGAVGNSHSSIGEFAYKSTCHFTCAAGFGLGQPAQIECTAQAQGWQQAIVCK 299
QY      60  VIOCEPLASPDGLGIMNCSH-PLASFSTSACTFICSEGTELIGKKTKTICSSSGIWSNPSP 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      300  AVCPAASQPKNGIYVFTHTSPTEGFTYKSSCAFCSEGFELRGSAQLACTSOGQMTQDEV 359
QY      119  ICQ 121
      ||
Db      360  SCQ 362
QY      119  ICQ 121

```


Db 360 SCO 362

RESULT 18
US-08-824-692-23
Sequence 23, Application US/08824692
Patent No. 6017703

GENERAL INFORMATION:

APPLICANT: Kinders, Robert J.

APPLICANT: Enfield, David L.

APPLICANT: Hass, G. Michael

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING

TITLE OF INVENTION: FOR OR MODULATING A TUMOR ASSOCIATED ANTIGEN

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/824,692

FILING DATE: 08-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 130001.404

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 240 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-824-692-23

Query Match 22.6%; Score 165.5; DB 3; Length 240;
Best Local Similarity 30.3%; Pred. No. 2.5e-10;
Matches 33; Conservative 24; Mismatches 47; Indels 5; Gaps 4;

QY 19 GN-EFSSQCAFSCSEGTNLTG-IEETTCGPFQNMSSPEPTCOVIOCEPLAPDLG--IM 74

Db 6 GNVEFGVAVYTCNEGYYLLGEINVRCDTDG-WTNDIPICEVYCLPVTAPENCKIVS 64

QY 75 NCSHPLASFSTACFTSCGTELGKTKTICSSGTSNSPTCOQL 123

Db 65 SAMPDREYHFGQAVRFVCSGKTEGDEMHCSDDGFWSKKEPKVEI 113

RESULT 19
5514582-32

Patent No. 5514582

APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.

TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID

IMMUNOGLOBULINS

NUMBER OF SEQUENCES: 43

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/185,670

FILING DATE: 21-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 986,931

FILING DATE: 08-DEC-1992

APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 32
LENGTH: 127
5514582-32

Query Match 22.3%; Score 163.5; DB 6; Length 127;
Best Local Similarity 28.8%; Pred. No. 1.9e-10;
Matches 36; Conservative 25; Mismatches 47; Indels 17; Gaps 6;

QY 14 CTH-----PFGNFS-----FSSQCAFSCSEGTNLTG-IEETTCGPFQNMSSPEPTCOV 60

Db 2 CGHPGDTPEGTFTLTGNAVFEYGVKAVYTCNEGYYLLGEINVRCDTDG-WTNDIPICEV 60

QY 61 IOCEPLAPDLG--IMNCSHPLASFSTACFTSCGTELGKTKTICSSGTSNSPTCOQL 118

Db 61 VKCLPYTAPENCKIVSSAMEPDREYHFGQAVRFVCSGKTEGDEMHCSDDGFWSKKEPKVEI 119

QY 119 ICQKL 123

Db 120 KCVEI 124

RESULT 20
5256642-6

Patent No. 5256642

APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,

MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN

; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.

TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT

RECEPTOR 1 (CRI) AND A THROMBOLYTIC AGENT, AND THE METHODS OF

USE THEREOF

NUMBER OF SEQUENCES: 30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/588,128

FILING DATE: 24-SEP-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 412,745

FILING DATE: 26-SEP-1989

APPLICATION NUMBER: 332,865

FILING DATE: 03-APR-1989

APPLICATION NUMBER: 176,532

FILING DATE: 01-APR-1988

SEQ ID NO: 6;

LENGTH: 1466
5256642-6

Query Match 22.3%; Score 163.5; DB 6; Length 1466;
Best Local Similarity 31.9%; Pred. No. 4e-09;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;

QY 1 CEPLAPELLGTDCTHPFNFSFSSQCAFSCSEGTNLTGIEETTCGPFQFN---WSSPEPT 57

Db 955 CGPEPEPFNGVY---HINTDQFGSTVNYSCNEGFRILGSPSTYCLVSNNTVWTKKAD 1011

QY 58 COVIOCEP---LSAPDLGIMNCSHPLASFSTACFTSCGTELGKTKTIC 108

Db 1012 CEIISCEPPTISNGE---YSNNFTSPHNGTVYVYQCHTGPDDGQLELVGERSICT 1067

QY 109 SS---GIMSNPSPTCOQLDK 125

Db 1068 SKDDQYGVWSSPPRCISTNK 1088

RESULT 21

5472939-6

Patent No. 5472939

```
APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138, 825
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588, 128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:6:
; LENGTH: 1466
5472939-6
```

```
Query Match          22.3%; Score 163.5; DB 6; Length 1466;
Best Local Similarity 31.9%; Pred. No. 4e-09;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;
```

```
QY 1 CEPLAPELLGTMDCTHPRGNFSSQCAFSCSEGTNLTGIEETTCGPGFN--WSSPEPT 57
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 955 CGPPEPENGWV---HINTDQFGSTVYVSCNCGFRIGSGSTTCVAGNNVTWDKKAPI 1011
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 CVOIOCEP---LSAPDGIAMNCSPHPLASFSTSACTFICSEGT-----ELIGKKKTICE 108
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1012 CEIISCEPPTISNGDF---YSNNRTSFHNGTVVYTCGHTGPDGDLFELVGERSIYCT 1067
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 109 SS-----GIWSNPSPIQCKLKD 125
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1068 SKDDGVGWSSPPRCISTNK 1088
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 22
5256642-5
; Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
; RECEPTOR 1 (CRI) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
; USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:5:
; LENGTH: 1537
5256642-5
```

```
Query Match          22.3%; Score 163.5; DB 6; Length 1537;
Best Local Similarity 31.9%; Pred. No. 4.2e-09;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;
```

```
QY 58 CVOIOCEP---LSAPDGIAMNCSPHPLASFSTSACTFICSEGT-----ELIGKKKTICE 108
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1012 CEIISCEPPTISNGDF---YSNNRTSFHNGTVVYTCGHTGPDGDLFELVGERSIYCT 1067
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 109 SS-----GIWSNPSPIQCKLKD 125
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1068 SKDDGVGWSSPPRCISTNK 1088
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 23
5472939-5
; Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138, 825
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588, 128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:5:
; LENGTH: 1537
5472939-5
```

```
Query Match          22.3%; Score 163.5; DB 6; Length 1537;
Best Local Similarity 31.9%; Pred. No. 4.2e-09;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;
```

```
QY 1 CEPLAPELLGTMDCTHPRGNFSSQCAFSCSEGTNLTGIEETTCGPGFN--WSSPEPT 57
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 955 CGPPEPENGWV---HINTDQFGSTVYVSCNCGFRIGSGSTTCVAGNNVTWDKKAPI 1011
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 CVOIOCEP---LSAPDGIAMNCSPHPLASFSTSACTFICSEGT-----ELIGKKKTICE 108
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1012 CEIISCEPPTISNGDF---YSNNRTSFHNGTVVYTCGHTGPDGDLFELVGERSIYCT 1067
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 109 SS-----GIWSNPSPIQCKLKD 125
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1068 SKDDGVGWSSPPRCISTNK 1088
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 24
5256642-10
; Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
; RECEPTOR 1 (CRI) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
; USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:10:
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:07 ; Search time 72.54 Seconds
(without alignments)
202.119 Million cell updates/sec

```

Title:      US-09-119-209-2_COPY_197_328
Perfect score: 733
Sequence:   1 CEPLAEPLGTMDCTHPFGN.....WSNPSPIQCKLKDKSFSMIKE 132

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

A.geneseq.032802.*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980. DAT: *
- 2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981. DAT: *
- 3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982. DAT: *
- 4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983. DAT: *
- 5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984. DAT: *
- 6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985. DAT: *
- 7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986. DAT: *
- 8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987. DAT: *
- 9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988. DAT: *
- 10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989. DAT: *
- 11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990. DAT: *
- 12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991. DAT: *
- 13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992. DAT: *
- 14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993. DAT: *
- 15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994. DAT: *
- 16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995. DAT: *
- 17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996. DAT: *
- 18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997. DAT: *
- 19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998. DAT: *
- 20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999. DAT: *
- 21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000. DAT: *
- 22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001. DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	733	100.0	369	17	AAR98125		Human Lymphocyte c
2	733	100.0	370	17	AAR98126		Human Lymphocyte c
3	733	100.0	370	17	AAR98127		Human Lymphocyte c
4	733	100.0	371	17	AAR98109		Human Lymphocyte c
5	733	100.0	371	17	AAR98110		Human Lymphocyte c
6	733	100.0	371	17	AAR98111		Human Lymphocyte c
7	733	100.0	371	17	AAR98112		Human Lymphocyte c
8	733	100.0	371	17	AAR98113		Human Lymphocyte c
9	733	100.0	371	17	AAR98114		Human Lymphocyte c
10	733	100.0	371	17	AAR98115		Human Lymphocyte c
11	733	100.0	371	17	AAR98116		Human Lymphocyte c

12	733	100.0	372	12	AAR12065	Human lymphocyte c
13	733	100.0	372	13	AAR24026	Sequence of human
14	733	100.0	372	14	AAR37960	Human lymphocyte h
15	733	100.0	372	14	AAR38908	HuLHR, Homo sapie
16	733	100.0	372	16	AAR76506	Human LHR, Homo s
17	733	100.0	372	16	AAR83050	Human LHR, Homo s
18	733	100.0	372	17	AAR89106	Human lymphocyte c
19	733	100.0	372	19	AAR37781	Homo sapiens lymph
20	733	100.0	372	20	AAW73264	Human lymphocyte h
21	733	100.0	374	17	AAR8131	Human lymphocyte c
22	733	100.0	374	17	AAR88132	Human lymphocyte c
23	732	99.9	371	17	AAR88122	Human lymphocyte c
24	731	99.7	371	17	AAR88124	Human lymphocyte c
25	730	99.6	371	17	AAR88119	Human lymphocyte c
26	729	99.5	371	17	AAR88117	Human lymphocyte c
27	729	99.5	371	17	AAR88121	Human lymphocyte c
28	728	99.3	372	13	AAR82802	Human lymphocyte h
29	727	99.2	371	17	AAR88118	Human lymphocyte c
30	727	99.2	371	17	AAR88120	Human lymphocyte c
31	727	99.2	371	17	AAR88123	Human lymphocyte c
32	727	99.2	372	22	AAB88834	Human lymphocyte c
33	722.5	98.6	372	17	AAR88133	Amino acid sequenc
34	721.5	98.4	374	17	AAR88134	Human lymphocyte c
35	721.5	98.4	374	17	AAR88135	Human lymphocyte c
36	719	98.1	363	13	AAR822551	T lymphocyte-speci
37	719	98.1	363	13	AAR81443	Human leu8 antigen
38	719	98.1	363	21	AAI96184	Human T-cell speci
39	719	98.1	385	13	AAR20815	T lymphocyte-speci
40	719	98.1	385	17	AAR91442	Human leu8 antigen
41	719	98.1	385	19	AAW80452	Human leu8 antigen
42	719	98.1	385	20	AAW68159	Human leu8 antigen
43	719	98.1	385	21	AAI96138	Human T-cell speci
44	719	98.1	405	22	AAU02447	Human T-lymphocyte
45	718.5	98.0	370	17	AAR88130	Human lymphocyte c

ALIGNMENTS	
RESULT 1	
AAR98125	
ID AAR98125	standard; Protein; 369 AA.
XX	
AC AAR98125;	
XX	
DT 01-NOV-1996	(first entry)
DE	
XX	
XX	Human lymphocyte cell surface glycoprotein (HULHR) variant.
KW Immunoglobulin;	transmembrane receptor; adhesion; targeting;
KW diagnosis;	therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator;	cell adhesion; graft rejection; inflammation;
KW metastasis.	
XX	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT Region	1..37
FT	/label= Signal region.
FT Domain	38..152
FT	/label= Lectin domain.
FT Domain	157..190
FT	/label= EGF domain.
FT Binding-site	194..255
FT	/label= Complement binding repeat 1.
FT Binding-site	256..314
FT	/label= Complement binding repeat 2.
FT Domain	330..352
FT	/label= Transmembrane domain.
FT Domain	353..369
FT	/label= Cytoplasmic domain.
XX	
PN US5514582-A.	

```
XX 07-MAY-1996.
PD
XX 23-FEB-1989; 89US-0315015.
PF
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant is a Gly96-Ile97 deletion mutant.
XX
XX Sequence 369 AA;
SQ
```

```
Query Match 100.0%; Score 733; DB 17; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CEPLAEPLCTMDCTHEFGNFSSOCAFSCSEGTNLGTIEETTCGPFQNMSSPEPTCOV 60
DB 194 cepleapeltgmdctbpfqnfssqcatcscsegtlgtieetctgpfqnmsspeptcy 253
QY 61 IOCEPIASAPLGLMNSHPLASFSFSACTFCTSEGTGLGKKTTCGSSGWTNSPSPIC 120
DB 254 iqcepisaplgimnshplastsfscactfctosegtelggkktlccssgwtlwnspspic 313
QY 121 OKLDKSFMSWIK 132
DB 314 qkldksfmswike 325
```

```
RESULT 2
AAR98126
ID AAR98126 standard; Protein; 370 AA.
XX
AC AAR98126;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HuLHR) variant.
```

```
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT Domain /label= Signal region.
FT Domain /label= 38..153
FT Domain /label= Lectin domain.
FT Domain /label= 158..191
FT Binding-site /label= EGF domain.
FT Binding-site /label= 195..256
FT Binding-site /label= Complement binding repeat 1.
FT Binding-site /label= 257..315
FT Domain /label= Complement binding repeat 2.
FT Domain /label= 331..353
FT Domain /label= Transmembrane domain.
FT Domain /label= 354..370
FT Domain /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant is an Asn136 deletion mutant.
XX
XX Sequence 370 AA;
SQ
```

```
Query Match 100.0%; Score 733; DB 17; Length 370;
```

Best Local Similarity	100.0%	Prod. No. 1.4e-63	Matches 132	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	CEPLEAPLGMDCIHPGNEFSFSQCAFSCSECTNLGTGIEETCGPGNMSSPEPTQY	60				
Db	195	cepleaplgmcdchpgrfnfsgfscqafscsegnlgtieetcgpfgnwsspeptcqy	254				
QY	61	IQCEPLSAPDGLIMNCSPHLASFSTSTACTFICSEGTLEIGKKTKTICSSSGIWSNPSPIC	120				
Db	255	iqceplsapdglimncshplasfststactficssegtlelqkkkticesssglwsnpspic	314				
QY	121	OKLDSFSMIKE	132				
Db	315	qklksfsmike	326				
RESULT 3							
ID	AA98127	standard; Protein: 370 AA.					
XX	AA98127:						
XX	01-NOV-1996	(first entry)					
XX	DE	Human lymphocyte cell surface glycoprotein (HLHR) variant.					
XX	KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;					
XX	KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;					
XX	KW	immunomodulator; cell adhesion; graft rejection; inflammation;					
XX	KW	metastasis.					
XX	OS	Homo sapiens.					
XX	Key	Location/Qualifiers					
XX	FT	1..37					
XX	FT	/label= Signal region.					
XX	FT	38..154					
XX	FT	/label= Lectin domain.					
XX	FT	159..191					
XX	FT	/label= EGF domain.					
XX	FT	195..256					
XX	FT	/label= Complement binding repeat 1.					
XX	FT	257..315					
XX	FT	/label= Complement binding repeat 2.					
XX	FT	331..353					
XX	FT	/label= Transmembrane domain.					
XX	FT	354..370					
XX	FT	/label= Cytoplasmic domain.					
XX	XX	US5514582-A.					
XX	PN	07-MAY-1996.					
XX	XX	23-FEB-1989;	89US-0315015.				
XX	XX	22-NOV-1989;	89US-0440625.				
XX	XX	23-FEB-1989;	89US-0315015.				
XX	XX	16-DEC-1991;	91US-0808122.				
XX	XX	08-DEC-1992;	92US-0986931.				
XX	XX	21-JAN-1994;	94US-0185670.				
XX	PA	(GETH) GENENTECH INC.					
XX	PI	Capon DJ, Lasky LA;					
XX	DR	WPI: 1996-238773/24.					
XX	XX	Nucleic acid encoding hybrid immunoglobulin comprising the ligand					
XX	XX	binding site of a receptor fused to Ig constant region - useful for					
XX	XX	diagnosis and treatment e.g. of inflammation					
XX	XX	Disclosure; Page 19; 41pp; English.					

CC	A hybrid immunoglobulin chain comprising the ligand binding site of
CC	a single transmembrane receptor without an active transmembrane
CC	region; fused at its C-terminus with the N-terminus of an
CC	immunoglobulin constant region. The receptor is not a member of the
CC	immunoglobulin super family, nor a multiple subunit polypeptide
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines
CC	the adhesion/targetting of a ligand binding pattern (LBP) with the
CC	effector functions of immunoglobulin and can bind to and/or activate
CC	more than one ligand. It can be used diagnostically for the in
CC	vitro assay of LBP and their targets, or therapeutically to deliver
CC	LBP such as toxins, enzymes, growth factors to particular cells.
CC	Typical applications are as antiviral, neuromodulating and
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC	The immunoglobulin component increases plasma half life and
CC	facilitates purification while deletion of the transmembrane region
CC	facilitates recovery, improves aqueous solubility and removes
CC	potentially immunogenic epitopes. Variants of the human lymphocyte
CC	cell surface glycoprotein described in AAR98106 are given in
CC	AAR98109-P98135. This variant is a Ser166 deletion mutant.
SQ	Sequence 370 AA;
OY	Query Match 100.0%; Score 733; DB 17; Length 370;
Db	Best Local Similarity 100.0%; Prid. No. 1,4e-63;
	Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CEPLAPELGTMDCNHPGNFSSQCAFSCSEGTNLGIERTTCGRGMNSPPETCOV 60
Db	195 ceplapeelgltndchpignfsfssqcafcscsegnltyleettcpgfnwsspctqv 254
OY	61 IQCEPLAPDGLIMNCSPHLASFSTTSACTFCSEGTLLIGKKTICSSGIMSNPPIIC 120
Db	255 iqceplsapdldgimncspilasfsttsactfcscsegtellgkktkcscsgtlwnppic 314
OY	121 QKLDKSFSMIKE 132
Db	315 qklksfsmike 326
RESULT 4	
AAR98109	
ID	AAR98109 standard; Protein: 371 AA.
AC	AAR98109;
XX	
DT	31-Oct-1996 (first entry)
XX	
DE	Human lymphocyte cell surface glycoprotein (HULDR) variant.
XX	
KW	Immunoglobulin; transmembrane receptor; adhesion; targetting;
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW	immunomodulator; cell adhesion; graft rejection; inflammation;
KW	metastasis.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Region 1..37
FT	/label= Signal region.
FT	Domain 38..154
FT	/label= lectin domain.
FT	Domain 159..192
FT	/label= EGF domain.
FT	Binding-site 196..257
FT	/label= Complement binding repeat 1.
FT	Binding-site 258..316
FT	/label= Complement binding repeat 2.
FT	Domain 332..354
FT	/label= Transmembrane domain.
FT	Domain 355..371
FT	/label= Cytoplasmic domain.

```

XX XX US514582-A.
XX XX
XX PD 07-MAY-1996.
XX XX
XX PF 23-FEB-1989; 89US-0315015.
XX XX
XX PR 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
XX PR 21-JAN-1994; 94US-0185670.
XX XX
XX PA (GETH ) GENENTECH INC.
XX PI Capon DJ, Lasky LA;
XX XX
XX DR WPI, 1996-238773/24.
XX XX
XX PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX PT binding site of a receptor fused to Ig constant region - useful for
XX PT diagnosis and treatment e.g. of inflammation
XX PS
XX PS Disclosure; Page 19; 41pp; English.
XX XX
XX CC A hybrid immunoglobulin chain comprising the ligand binding site of
XX CC a single transmembrane receptor without an active transmembrane
XX CC region; fused at its C-terminus with the N-terminus of an
XX CC immunoglobulin constant region. The receptor is not a member of the
XX CC immunoglobulin super family, nor a multiple subunit polypeptide
XX CC encoded by discrete genes. The hybrid immunoglobulin chain combines
XX CC the adhesion/targeting of a ligand binding partner (LBP) with the
XX CC effector functions of immunoglobulin and can bind to and/or activate
XX CC more than one ligand. It can be used diagnostically for the in
XX CC vitro assay of LBP and their targets; or therapeutically to deliver
XX CC LBP such as toxins, enzymes, growth factors to particular cells.
XX CC Typical applications are as antiviral, neuromodulating and
XX CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX CC The immunoglobulin component increases plasma half life and
XX CC facilitates purification while deletion of the transmembrane region
XX CC facilitates recovery, improves aqueous solubility and removes
XX CC potentially immunogenic epitopes. Variants of the human lymphocyte
XX CC cell surface glycoprotein described in AAR98106 are given in
XX CC AAR98109-R98135. This variant contains a Arg58-Asp59; Lys-Glu
XX CC substitution.
XX CC
XX CC Sequence 371 AA;
XX SQ

```

```

DT 01-NOV-1996 (first entry)
XX XX
XX DE Human lymphocyte cell surface glycoprotein (HLHR) variant.
XX XX
XX KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX KW immunomodulator; cell adhesion; graft rejection; inflammation;
XX KW metastasis.
XX XX
XX OS Homo sapiens.
XX XX
XX FH
XX FH Key
XX FH Region
XX FH
XX FH Domain
XX FH
XX FH Domain
XX FH
XX FH Binding-site
XX FH
XX FH Binding-site
XX FH
XX FH Binding-site
XX FH
XX FH Domain
XX FH
XX FH Domain
XX FH
XX PN US514582-A.
XX XX
XX XX 07-MAY-1996.
XX PF
XX PF 23-FEB-1989; 89US-0315015.
XX XX
XX PR 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
XX PR 21-JAN-1994; 94US-0185670.
XX XX
XX PA (GETH ) GENENTECH INC.
XX PI Capon DJ, Lasky LA;
XX XX
XX DR WPI, 1996-238773/24.
XX XX
XX PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX PT binding site of a receptor fused to Ig constant region - useful for
XX PT diagnosis and treatment e.g. of inflammation
XX PS
XX PS Disclosure; Page 19; 41pp; English.
XX XX
XX CC A hybrid immunoglobulin chain comprising the ligand binding site of
XX CC a single transmembrane receptor without an active transmembrane
XX CC region; fused at its C-terminus with the N-terminus of an
XX CC immunoglobulin constant region. The receptor is not a member of the
XX CC immunoglobulin super family, nor a multiple subunit polypeptide
XX CC encoded by discrete genes. The hybrid immunoglobulin chain combines
XX CC the adhesion/targeting of a ligand binding partner (LBP) with the
XX CC effector functions of immunoglobulin and can bind to and/or activate
XX CC more than one ligand. It can be used diagnostically for the in
XX CC vitro assay of LBP and their targets; or therapeutically to deliver
XX CC LBP such as toxins, enzymes, growth factors to particular cells.
XX CC Typical applications are as antiviral, neuromodulating and
XX CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX CC The immunoglobulin component increases plasma half life and
XX CC facilitates purification while deletion of the transmembrane region
XX CC facilitates recovery, improves aqueous solubility and removes
XX CC potentially immunogenic epitopes. Variants of the human lymphocyte
XX CC cell surface glycoprotein described in AAR98106 are given in
XX CC AAR98109-R98135. This variant contains an Ala71ser substitution.
XX CC
XX CC Sequence 371 AA;
XX SQ

```



```

FT FT /label= Transmembrane domain.
FT Domain 355...371
XX /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 16-DEC-1991; 91US-0808122.
XX 08-DEC-1992; 92US-0986931.
XX 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to Ig constant region - useful for
XX diagnosis and treatment e.g. of inflammation
XX
XX PT diagnosis and treatment e.g. of inflammation
XX
XX PS Disclosure; Page 19; 41pp; English.
XX
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
XX immunoglobulin constant region. The receptor is not a member of the
XX immunoglobulin super family, nor a multiple subunit polypeptide
XX encoded by discrete genes. The hybrid immunoglobulin chain combines
XX the adhesion/targeting of a ligand binding partner (LBP) with the
XX effector functions of immunoglobulin and can bind to and/or activate
XX more than one ligand. It can be used diagnostically for the in
XX vitro assay of LBP and their targets; or therapeutically to deliver
XX LBP such as toxins, enzymes, growth factors to particular cells.
XX Typical applications are as antiviral, neuromodulating and
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX The immunoglobulin component increases plasma half life and
XX facilitates purification while deletion of the transmembrane region
XX facilitates recovery, improves aqueous solubility and removes
XX potentially immunogenic epitopes. Variants of the human lymphocyte
XX cell surface glycoprotein described in AAR98106 are given in
XX AAR98109-R98135. This variant contains an Asp116Glu substitution.
XX
XX SQ Sequence 371 AA:

```

```

AC AAR98113;
XX
XX 01-NOV-1996 (first entry)
XX
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX immunomodulator; cell adhesion; graft rejection; inflammation;
XX metastasis.
XX
XX OS Homo sapiens.
XX
XX FH Key
XX Region 1..37
XX Domain /label= Signal region.
XX Domain 38..154
XX Domain /label= Lectin domain.
XX Domain 159..192
XX Binding-site /label= EGF domain.
XX Binding-site 196..257
XX Binding-site /label= Complement binding repeat 1.
XX Binding-site 258..316
XX Domain /label= Complement binding repeat 2.
XX Domain 332..354
XX Domain /label= Transmembrane domain.
XX Domain 355..371
XX /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 16-DEC-1991; 91US-0808122.
XX 08-DEC-1992; 92US-0986931.
XX 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to Ig constant region - useful for
XX diagnosis and treatment e.g. of inflammation
XX
XX PS Disclosure; Page 19; 41pp; English.
XX
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
XX immunoglobulin constant region. The receptor is not a member of the
XX immunoglobulin super family, nor a multiple subunit polypeptide
XX encoded by discrete genes. The hybrid immunoglobulin chain combines
XX the adhesion/targeting of a ligand binding partner (LBP) with the
XX effector functions of immunoglobulin and can bind to and/or activate
XX more than one ligand. It can be used diagnostically for the in
XX vitro assay of LBP and their targets; or therapeutically to deliver
XX LBP such as toxins, enzymes, growth factors to particular cells.
XX Typical applications are as antiviral, neuromodulating and
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX The immunoglobulin component increases plasma half life and
XX facilitates purification while deletion of the transmembrane region
XX facilitates recovery, improves aqueous solubility and removes
XX potentially immunogenic epitopes. Variants of the human lymphocyte
XX cell surface glycoprotein described in AAR98106 are given in
XX AAR98109-R98135. This variant contains a Leu150Val substitution.
XX

```

```
XX SQ Sequence 371 AA:
Query Match 100.0%; Score 733; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEPLAEPLCTMDCTHPEGNFSSQCAFSCSEGTNLGTIEETTCGPFGMWSSPEPTCOV 60
DB 196 CEPLAEPLCTMDCTHPEGNFSSQCAFSCSEGTNLGTIEETTCGPFGMWSSPEPTCOV 255
QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTELIGKKTTCESGIMSNPSPIC 120
DB 256 IGCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTELIGKKTTCESGIMSNPSPIC 315
QY 121 QKLDKSFMSIKE 132
DB 316 QKLDKSFMSIKE 327
RESULT 9
AAR98114
ID AAR98114 standard; Protein; 371 AA.
AC AAR98114;
XX
DT 01-NOV-1996 (first entry)
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KM Immunoglobulin; transmembrane receptor; adhesion; targeting;
KM diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KM immunomodulator; cell adhesion; graft rejection; inflammation;
KM metastasis.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 1..37 /label= Signal region.
FT Domain 38..154 /label= Lectin domain.
FT Domain 159..192 /label= EGF domain.
FT Binding-site 196..257 /label= Complement binding repeat 1.
FT Binding-site 258..316 /label= Complement binding repeat 2.
FT Domain 332..354 /label= Transmembrane domain.
FT Domain 355..371 /label= Cytoplasmic domain.
XX
PN US5514582-A.
XX
PD 07-MAY-1996.
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
```

```
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an H1S168Gln substitution.
XX
SQ Sequence 371 AA:
Query Match 100.0%; Score 733; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEPLAEPLCTMDCTHPEGNFSSQCAFSCSEGTNLGTIEETTCGPFGMWSSPEPTCOV 60
DB 196 CEPLAEPLCTMDCTHPEGNFSSQCAFSCSEGTNLGTIEETTCGPFGMWSSPEPTCOV 255
QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTELIGKKTTCESGIMSNPSPIC 120
DB 256 IGCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTELIGKKTTCESGIMSNPSPIC 315
QY 121 QKLDKSFMSIKE 132
DB 316 QKLDKSFMSIKE 327
RESULT 10
AAR98115
ID AAR98115 standard; Protein; 371 AA.
AC AAR98115;
XX
DT 01-NOV-1996 (first entry)
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KM Immunoglobulin; transmembrane receptor; adhesion; targeting;
KM diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KM immunomodulator; cell adhesion; graft rejection; inflammation;
KM metastasis.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 1..37 /label= Signal region.
FT Domain 38..154 /label= Lectin domain.
FT Domain 159..192 /label= EGF domain.
FT Binding-site 196..257 /label= Complement binding repeat 1.
FT Binding-site 258..316 /label= Complement binding repeat 2.
```

FT		/label= Complement binding repeat 2.
PT	Domain	332..354
FT		/label= Transmembrane domain.
PT	Domain	355..371
FT		/label= Cytoplasmic domain.
XX		
PN	US5514582-A.	
XX		
PD	07-MAY-1996.	
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GENTH) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI; 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation	
XX		
PS	Disclosure; Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region; fused at its C-terminus with the N-terminus of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targeting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the in vitro assay of LBP and their targets; or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particular cells. Typical applications are as antiviral, neuromodulating and immunomodulating agents, or as modulators of cell adhesion (e.g. in treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery. Improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR98106 are given in AAR8109-R88135. This variant contains an Ile174Leu substitution.	
CC		
XX	Sequence ` 371 AA:	
Query Match	100.0%: Score 733; DB 17; Length 371;	
Best Local Similarity	100.0%; Pred. NO. 1,4e-63;	
Matches 132; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 CEPLEAPELGTMDCHHPFGNFSFSSOCAFCSEEGTNLTGIETTGCPFMNMSSPEPTCOY 60 	
Db	196 cepleapeelgtmdchhpfgnfsfssqcafcseegnlgielttcgpfmfwsspptcyy 255 	
OY	61 IOCEPLSNAPDLGINMCSHRPLASFSTTSACFTFCSGGTELIGKKKTTICSSSGIWSNPSPIC 120 	
Db	256 iqceplsapdldginmcshrpilasfsttsactffcsggteligkkkticesssgilwspnpic 315 	
OY	121 OKLDKSFMSMIKE 132 	
Db	316 qkldksfmike 327	
RESULT	11	
AAR981116		

ID	AA08116	standard; Protein; 371 AA.
XX	AA08116;	
AC	AA08116;	
XX		
DT	01-NOV-1996	(first entry)
XX		
DE	Human lymphocyte cell surface glycoprotein (HLHR) variant.	
XX		
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;	
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KW	immunomodulator; cell adhesion; graft rejection; inflammation;	
KW	metastasis.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= EGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1.
FT	Binding-site	258..316
FT		/label= Complement binding repeat 2.
FT	Domain	332..354
FT		/label= Transmembrane domain.
FT	Domain	355..371
FT		/label= Cytoplasmic domain.
XX		
XX		
XX	US514582-A.	
PN		
XX		
PD	07-MAY-1996.	
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
XX		
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GETH) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI; 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
XX		
PS	Disclosure; Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particularly to cell.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.).	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while detection of the transmembrane region	
CC	facilitates recovery. Improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	

CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-R98135. This variant contains an Asn181Gln substitution.
 XX
 SQ Sequence 371 AA;

Query Match 100.0%; Score 733; DB 17; Length 371;
 Best Local Similarity 100.0%; Pred. No. 1.4e-63;
 Matches 132: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTQV 60
 |||||||
 DB 196 ceplapeligmdcthpngfnfsfscsfscsegtnlgieetctgpgnwspeptcqy 255
 QY 61 IOCEPLASPDGLGIMNCSPHPLASFSTSACTFICSGTELLIGKRTICSSGIGWSPSPIC 120
 |||||||
 DB 256 iqceplaspdglgimncshplasfstsaactficsgtellgkktlcicssgigwnspspic 315
 QY 121 QKDKSFSMIKE 132
 |||||||
 DB 316 qkdkfsfmike 327

RESULT 12

AAR12469
 ID AAR12469 standard; Protein; 372 AA.

AC AAR12469;

DT 09-SEP-1991 (first entry)

XX Human lymphocyte cell surface glycoprotein (LHR).

XX
 KW Ligand binding partner: lbp; stable plasma protein; spp; antiviral;
 immunomodulatory; neuromodulatory; receptor mediated abnormality.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 20..32
 FT /label= signal peptide
 FT Protein 38..372
 FT /label= probable mature protein

PN MO9108298-A.

XX 13-JUN-1991.

XX 21-NOV-1990; 90WO-US06849.

XX 22-NOV-1989; 89US-0444625.

PA (GETH) GENENTECH INC.

PI Capon DJ, Lasky LA;

XX WPI: 1991-199202/26.

DR N-PSDB: AAQ12118.

XX New hybrid immunoglobulin(s) - for use as diagnostic reagents for
 PT ligand binding molecules and to treat organ and graft rejection
 PT and inflammation.

XX Disclosure: Fig 1; 67pp; English.

XX The gene product may be used as a ligand binding partner in combina-
 CC tion with a stable plasma protein eg. IgG1-IgG4, IgA, IgE, IgD or IgM.
 CC The fusion product is joined by N- or C-terminal groups, preferably
 CC the N-terminal of the fc region of the spp is linked to the C-terminal
 CC of lbp. They may be used to provide antiviral, immunomodulatory
 CC and neuromodulatory treatment as well as in treatment of receptor
 CC mediated abnormalities.

SQ Sequence 372 AA;

Query Match 100.0%; Score 733; DB 12; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.4e-63;
 Matches 132: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTQV 60
 |||||||
 DB 197 ceplapeligmdcthpngfnfsfscsfscsegtnlgieetctgpgnwspeptcqy 256
 QY 61 IOCEPLASPDGLGIMNCSPHPLASFSTSACTFICSGTELLIGKRTICSSGIGWSPSPIC 120
 |||||||
 DB 257 iqceplaspdglgimncshplasfstsaactficsgtellgkktlcicssgigwnspspic 316
 QY 121 QKDKSFSMIKE 132
 |||||||
 DB 317 qkdkfsfmike 328

RESULT 13

AAR24026
 ID AAR24026 standard; Protein; 372 AA.

AC AAR24026;

DT 22-NOV-1992 (first entry)

XX Sequence of human lymphocyte cell surface glycoprotein
 DE (HLHR).

XX Lymphocyte cell surface glycoprotein; ligand binding protein.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 1..19
 FT Peptide 20..38
 FT /label= signal
 FT Modified-site 60..62
 FT /label= potential N-linked glycosylation site
 FT Modified-site 104..106
 FT /label= see above
 FT Modified-site 177..179
 FT /label= see above
 FT Modified-site 216..218
 FT /label= see above
 FT Modified-site 232..234
 FT /label= see above
 FT Modified-site 271..273
 FT /label= see above
 FT Modified-site 311..313
 FT /label= see above
 FT Region 333..355
 FT /label= stop transfer sequence

PN US5116964-A.

XX 26-MAY-1992.

XX 22-NOV-1989; 89US-0440625.

XX 23-FEB-1989; 88US-0315015.

PR 22-NOV-1989; 89US-0440625.

PA (GETH) GENENTECH INC.

PI Capon DJ, Lasky LA;

XX WPI: 1992-199589/24.

DR N-PSDB: AAQ24987.

XX Nucleic acid encoding polypeptide fusions - comprising ligand

PT binding partner protein and immunoglobulin chain, for use in
PT diagnosis and therapy
XX
PS Disclosure; Fig 1-1 - 1-3; 43pp; English.
XX
CC LHR mediates the binding of lymphocytes to the endothelium of
CC lymphoid tissue. Full length cDNA clones and DNA encoding the human
CC and the murine LHR (HLHR and MLHR, respectively) have been
CC identified and isolated (see AA024987 and AA024988). LHR is a
CC glycoprotein which contains the following protein domains: a signal
CC sequence, a carbohydrate binding domain, and epidermal growth
CC factor-like (egf) domain, at least one and preferably two complement
CC binding domain repeat, a transmembrane binding domain (TMD), and a
CC charged intracellular or cytoplasmic domain. LHR is used as the
CC ligand-binding partner in fusion polypeptides with an immunoglobulin,
CC for use in diagnosis and therapy.
XX
SQ Sequence 372 AA;

Query Match 100.0%; Score 733; DB 13; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEPLAPELLGTMDCTHPGNSFSOCAFSCSGTNTLTGIEETTCGPGNWSPEPTCOY 60
|||||
DB 197 CEPLAPELLGTMDCTHPGNSFSOCAFSCSGTNTLTGIEETTCGPGNWSPEPTCOY 256
|||||
QY 61 IOCEPLSAPDLGIMNCSHPLASFSFTSACTFTICSEGTETLIGKKKTICSSGIMNSPPTC 120
|||||
DB 257 IGCEPLSAPDLGIMNCSHPLASFTSACTFTICSEGTETLIGKKKTICSSGIMNSPPTC 316
|||||
QY 121 OKDKSFSMIKE 132
|||||
DB 317 qkldksfsmike 328

RESULT 14
AA037960
ID AA037960 standard; protein; 372 AA.
XX
AC AA037960;
XX
DT 08-OCT-1993 (first entry)
XX
DE Human Lymphocyte Homing Receptor.
XX
KW HLHR; lymphocyte binding inhibition; lymphoma metastasis;
KW transplant rejection; inflammation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..38
FT /label= signal_sequence
FT Protein 39..372
FT /note= "Trp39 is probable N-terminus of mature LHR"
FT Domain 39..155
FT /label= Lectin_domain
FT Modified-site 60..62
FT /note= "potential N-glycosylation site"
FT Modified-site 104..106
FT /note= "potential N-glycosylation site"
FT Domain 160..193
FT /label= BGF_domain
FT Modified-site 177..179
FT /note= "potential N-glycosylation site"
FT Region 197..258
FT /label= Complement_Binding_Repeat_1
FT Modified-site 216..218
FT /note= "potential N-glycosylation site"
FT Modified-site 232..234
FT /note= "potential N-glycosylation site"

FT Modified-site 246..248
FT /note= "potential N-glycosylation site"
FT Region 259..317
FT /label= Complement_Binding_Repeat_2
FT Modified-site 271..273
FT /note= "potential N-glycosylation site"
FT Modified-site 311..313
FT /note= "potential N-glycosylation site"
FT Domain 333..355
FT /label= Transmembrane_Domain
FT /note= "stop transfer sequence"
FT Domain 356..372
FT /label= Cytoplasmic_Domain
XX
PN US5216131-A.
XX
PD 01-JUN-1993.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 23-FEB-1989; 89US-0315015.
XX
PR 23-FEB-1989; 89US-0315015.
PR 31-OCT-1991; 91US-0786149.
XX
PA (GETH) GENENTECH INC.
XX
XX Lasky LA, Rosen SD, Singer MS, Stachel SE, Yednock TA;
XX
XX WPI: 1993-188588/23.
XX
DR N-PSDB; AA043154.
XX
XX
XX Human and murine lymphocyte homing receptors to treat graft
PT rejection and inflammation - comprise carbohydrate binding,
PT epidermal growth factor and complement binding domains
XX
PS Claim 1; Fig 1 and Fig 3; 32pp; English.
XX
XX A human peripheral blood lymphocyte cDNA library in lambda gt10 was
CC screened with a 2.2kb EcoRI insert of the murine Mel14 antigen clone
CC (i.e. a murine LHR sequence). The largest EcoRI insert (2.2kb) was
CC isolated and sequenced. The ORF codes for 372 amino acids with a mol.
CC wt. of approximately 42,200. Comparison of the HLHR amino acid
CC sequence with the murine LHR sequence (AA037961) showed a high degree
CC of amino acid conservation in each of the LHR domains, e.g. 96% in
CC the transmembrane domain and 83% in the carbohydrate binding domain.
CC The LHRs could be used to compete with the normal binding of
CC lymphocytes to lymphoid tissue to treat inflammation or graft
CC rejection. They could also be used to control lymphoma metastasis
CC and to treat conditions involving lymphocyte accumulation.
XX
SQ Sequence 372 AA;

Query Match 100.0%; Score 733; DB 14; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEPLAPELLGTMDCTHPGNSFSOCAFSCSGTNTLTGIEETTCGPGNWSPEPTCOY 60
|||||
DB 197 CEPLAPELLGTMDCTHPGNSFSOCAFSCSGTNTLTGIEETTCGPGNWSPEPTCOY 256
|||||
QY 61 IOCEPLSAPDLGIMNCSHPLASFSFTSACTFTICSEGTETLIGKKKTICSSGIMNSPPTC 120
|||||
DB 257 IGCEPLSAPDLGIMNCSHPLASFTSACTFTICSEGTETLIGKKKTICSSGIMNSPPTC 316
|||||
QY 121 OKDKSFSMIKE 132
|||||
DB 317 qkldksfsmike 328

RESULT 15
AA038908
ID AA038908 standard; protein; 372 AA.
XX

AC AAR38908;
 XX
 DT 11-JAN-1994 (first entry)
 XX
 DE HuLHR.
 XX
 KW Human; murine; lymphocyte; cell surface glycoprotein; homing receptor;
 KW LHR; endothelium; lymphoid tissue; signal; domain; complement binding;
 KW carbohydrate binding; epidermal growth factor-like; egf; intracellular;
 KW transmembrane binding; cytoplasmic; ligand binding partner protein;
 KW TMD; LBP.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /note="Signal peptide"
 FT 20..32
 FT 39..372
 FT /note="Mature protein"
 FT 39..155
 FT /note="Lectin domain"
 FT 160..193
 FT /note="egf domain"
 FT 197..317
 FT /note="Complement factor binding domain"
 FT 333..355
 FT /note="Transmembrane binding domain"
 FT 356..372
 FT /note="Cytoplasmic domain"
 XX
 PN US5225538-A.
 XX
 PD 06-JUL-1993.
 XX
 PF 23-FEB-1989; 8905-0315015.
 XX
 PR 23-FEB-1989; 8905-0315015.
 PR 22-NOV-1989; 8905-0440625.
 PR 16-DEC-1991; 9105-0808122.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 DR WPI: 1993-226664/28.
 DR N-PSDB; AA044243.
 XX
 PT New lymphocyte homing receptor immunoglobulin fusion
 PT polypeptide(s) - used to inhibit binding of lymphocytes in
 PT therapeutic and diagnostic uses
 XX
 PS Disclosure; Fig 1; 44pp; English.
 XX
 CC The sequences given in AAR38908-09 represent human and murine lymphocyte
 CC cell surface glycoprotein (LHR) respectively. These proteins mediate
 CC the binding of lymphocytes to the endothelium of lymphoid tissue. LHR
 CC is a glycoprotein which contains a signal domain, a carbohydrate
 CC binding domain, an epidermal growth factor-like (egf) domain, at least
 CC one complement binding domain repeat, a transmembrane binding domain
 CC (TMD) and a charged intracellular or cytoplasmic domain. The murine
 CC and human amino acid sequences show a high degree of overall homology
 CC (83%), however degrees of homology between the various domains is
 CC variable. These proteins may be fused to a ligand binding partner
 CC protein (LBP) which causes an increase in the half life of the LHR.
 CC The fusions may be used therapeutically to compete with the normal
 CC binding of lymphocytes to lymphoid tissue. They may be used in organ
 CC or graft rejection and for the treatment of inflammation.
 XX
 SO Sequence 372 AA;

Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEPLAPRLGTMDCITPRGNSFSQCAFSCSSEGNLTGIRETTGCPRGMSNPPTQOV 60
 DB 197 cepleaprlgimdcitprgnfsqcafsccsegnltgiretcgptgnwspptqv 256
 QY 61 IQCEPLAPDGLGIMNCSHPLASFSTSACTFICSEGTLEICKKTKTICSSGIMWSPIC 120
 DB 257 iqceplaspdlgimncshplasfstsaactficegtleigkkkticessgimwspic 316
 QY 121 QKLDKFSFMKE 132
 DB 317 qkldksfmike 328
 RESULT 16
 AAR76506
 ID AAR76506 standard; Protein: 372 AA.
 XX
 AC AAR76506;
 XX
 DT 01-DEC-1995 (first entry)
 XX
 DE Human LHR.
 XX
 KW Lymphocyte homing receptor; lymphocyte cell surface glycoprotein;
 KW LHR; ligand binding partner; immunoglobulin; constant region;
 KW antibody engineering; immunomodulator.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /note="Signal peptide"
 FT /note="hydrophobic domain, may act as signal
 FT for insertion into the endoplasmic
 FT reticulum lumen"
 FT 39..372
 FT /label="Mat-protein"
 FT 39..334
 FT /note="putative extracellular domain"
 FT 60..62
 FT /label="N-glycosylation-site"
 FT 104..106
 FT /label="N-glycosylation-site"
 FT 177..179
 FT /label="N-glycosylation-site"
 FT 216..218
 FT /label="N-glycosylation-site"
 FT 232..234
 FT /label="N-glycosylation-site"
 FT 246..248
 FT /label="N-glycosylation-site"
 FT 271..273
 FT /label="N-glycosylation-site"
 FT 311..313
 FT /label="N-glycosylation-site"
 FT 335..357
 FT /note="putative stop transfer or membrane anchor
 FT domain"
 FT 358..372
 FT /note="putative intracellular region"
 XX
 PN US5428130-A.
 XX
 PD 27-JUN-1995.
 XX
 PF 23-FEB-1989; 8905-0315015.
 XX
 PR 22-NOV-1989; 8905-0440625.
 PR 23-FEB-1989; 8905-0315015.
 PR 16-DEC-1991; 9105-0808122.
 PR 08-DEC-1992; 9205-0986931.

```

XX (GETH ) GENENTECH INC.
PA
XX
XX Capon DJ, Lasky LA;
PI
XX
XX WPI: 1995-240086/31.
DR
XX N-PSDB: AAQ92802.
PT
XX New hybrid ligand binding partner molecules - fused to immunoglobulin
PT constant region sequences to increase stability and in vivo plasma
PT half-life
XX
XX
XX Disclosure; Fig.1a-1c; 40pp; English.
PS
XX
XX A murine Mel 14 antigen cDNA clone was used to screen a lambda gt10
CC cDNA library derived from human peripheral blood lymphocyte mRNA
CC obtd. from primary cells. A cDNA clone encoding LHR was isolated.
XX
XX
XX Sequence 372 AA;
SQ

Query Match 100.0%; Score 733; DB 16; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAPELLGMDCTHPFGNPSFSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
DB 197 ceplaeplgmdcthpfgnfsfsqcafscegnltlgieettcgpgnsspeptcqv 256
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 120
DB 257 igceplsapdlgimncshplastsfstsactficsgegtelgkkticssgimnspspic 316
OY 121 QKDKSFSMIKE 132
DB 317 qkdkfsfmike 328

RESULT 17
AAR83050
ID AAR83050 standard; Protein; 372 AA.
XX
XX AAR83050;
AC
XX
XX 31-JAN-1996 (first entry)
DT
XX
XX Human LHR.
DE
XX
XX Lymphocyte cell surface glycoprotein; LHR; transmembrane receptor;
KW immunoglobulin; Igg: constant region; receptor-mediated disease;
KW vector; plasma-life.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 20..32
FT /label= Sig-peptide
FT 39..155
FT /label= Carbohydrate_binding_domain
FT 160..193
FT /label= Epidermal_growth_factor_domain
FT 197..317
FT /label= Complement_factor_binding_domain
FT 333..355
FT /label= Transmembrane_binding_domain
FT 356..372
FT /label= Cytoplasmic_domain
FT /note= "potential stop transfer sequence"
FT Modified-site 60..62
FT /label= N-glycosylation_site
FT Modified-site 104..106
FT /label= N-glycosylation_site
FT Modified-site 177..179

```

```

FT /label= N-glycosylation_site
FT 216..218
FT Modified-site /label= N-glycosylation_site
FT 232..234
FT Modified-site /label= N-glycosylation_site
FT 248..248
FT Modified-site /label= N-glycosylation_site
FT 271..273
FT Modified-site /label= N-glycosylation_site
FT 311..313
FT /label= N-glycosylation_site
XX
XX US5455165-A.
XX
XX 03-OCT-1995.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 16-DEC-1991; 91US-0808122.
XX 08-DEC-1992; 92US-0986931.
XX 21-JAN-1994; 94US-0185669.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
PI
XX
XX WPI: 1995-350776/45.
DR N-PSDB: AAT05869.
PT
PT Expression vector encoding fusion protein to increase plasma life -
PT comprises receptor ligand binding site and Ig constant region, for
PT treatment of receptor mediated disease
XX
XX Disclosure; Fig 1; 42pp; English.
XX
XX A mouse LHR (lymphocyte cell surface glycoprotein) cDNA clone was used
CC to screen an oligo-dt primed lambda gt10 cDNA library derived from
CC human peripheral blood lymphocyte mRNA obtd. from primary cells. A
CC 2.2 kb clone (sequence given in AAT05869) was isolated that encoded the
CC human LHR protein (AAR83050). LHR-igg hybrids were constructed for use
CC in the targeting of therapeutic moieties to lymphoid tissue.
XX
XX Sequence 372 AA;
SQ

Query Match 100.0%; Score 733; DB 16; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAPELLGMDCTHPFGNPSFSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
DB 197 ceplaeplgmdcthpfgnfsfsqcafscegnltlgieettcgpgnsspeptcqv 256
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 120
DB 257 igceplsapdlgimncshplastsfstsactficsgegtelgkkticssgimnspspic 316
OY 121 QKDKSFSMIKE 132
DB 317 qkdkfsfmike 328

RESULT 18
AAR98106
ID AAR98106 standard; Protein; 372 AA.
XX
XX AAR98106;
AC
XX
XX 31-OCT-1996 (first entry)
DT
XX
XX Human lymphocyte cell surface glycoprotein (HuLHR).
DE

```


XX Prevention of lymphocyte attachment to endothelial cells - using
 PT chimeric molecule comprising lymphocyte homing receptor and
 PT immunoglobulin constant region
 XX
 PS Disclosure; Fig 1; 43pp; English.
 XX

CC The sequence is that of a human lymphocyte homing receptor
 CC (LHR) which may be used in the construction of a chimeric molecule
 CC comprising an LHR fused at its C terminus to the N terminus of an
 CC immunoglobulin constant region. This can be used for the prevention
 CC of lymphocyte attachment to endothelial cells. Such a method may
 CC be used for preventing organ or graft rejection. For treating
 CC inflammatory disorders, e.g. rheumatoid arthritis or other
 CC autoimmune diseases, for controlling lymphoma metastasis and
 CC for treating conditions in which there is an accumulation of
 CC lymphocytes.
 CC
 CC Sequence 372 AA;

Query Match 100.0%; Score 733; DB 19; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.4e-63;
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAEPLGTMDCITHPFNGNFSFSSQCAFSCSEGTNLGIETTCGPFQWSSPEPTQY 60
 Db 197 ceplaeplgltmdcthpfnfsgafscsegtnlgieltcpgfnwsspeptcqv 256
 OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELKRRKKTICSSGIMNSPSPIC 120
 Db 257 iqceplsapdlgimncshplasfstsactficsegteltgkrrkticessgimwnpspic 316
 OY 121 QKDKSFSMIKE 132
 Db 317 qkdkfsfmike 328

RESULT 20
 AAW73264
 ID AAW73264 standard; Protein; 372 AA.
 AC AAW73264;
 XX
 DT 02-FEB-1999 (first entry)
 XX

DE Human lymphocyte homing receptor.
 XX
 KW Lymphocyte homing receptor; LHR; lymphocyte cell-surface glycoprotein;
 KW arthritis; autoimmune disease; lymphoma metastasis;
 KW lymphocyte accumulation; human.
 XX

OS Homo sapiens.
 XX
 PN US5840844-A.
 XX
 PD 24-NOV-1998.
 XX
 PF 10-AUG-1995; 95US-0513278.
 XX
 PR 23-FEB-1989; 89US-0315015.
 PR 31-OCT-1991; 91US-0786149.
 PR 06-MAY-1993; 93US-0059029.
 PR 10-AUG-1995; 95US-0513278.
 XX

PA (GENTH) GENENTECH INC.
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Lasky LA, Rosen SD, Singer MS, Stachel SE;
 XX
 DR WPI; 1999-034122/03.
 DR N-PDB; AAY08321.

XX Lymphocyte homing receptor polypeptides - useful for inhibiting
 PT lymphocyte binding to lymphoid endothelium
 PT
 XX
 PS Claim 1; Fig 1; 33pp; English.
 XX

CC This sequence is the human lymphocyte homing receptor (LHR) of the
 CC invention. LHR is a lymphocyte cell-surface glycoprotein that mediates
 CC the binding of lymphocytes to the endothelium of lymphoid tissue. Soluble
 CC LHR polypeptides, lacking signal peptide (amino acids 1-38),
 CC transmembrane domain (amino acids 333-355) and cytoplasmic domain (amino
 CC acids 356-372), can be used therapeutically to compete with the normal
 CC binding of lymphocytes to lymphoid tissue and are especially useful for
 CC organ or graft rejection treatment protocols, for treating inflammations
 CC such as arthritis and other autoimmune diseases, for control of lymphoma
 CC metastasis and for treating conditions involving lymphocyte accumulation.
 CC LHR polypeptides can also be used in assays for LHR, anti-LHR antibodies
 CC or competitive inhibitors of LHR activity and for purifying anti-LHR
 CC antibodies, and as immunogens for raising anti-LHR antibodies.
 CC
 CC Sequence 372 AA;

Query Match 100.0%; Score 733; DB 20; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.4e-63;
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAEPLGTMDCITHPFNGNFSFSSQCAFSCSEGTNLGIETTCGPFQWSSPEPTQY 60
 Db 197 ceplaeplgltmdcthpfnfsgafscsegtnlgieltcpgfnwsspeptcqv 256
 OY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTTELKRRKKTICSSGIMNSPSPIC 120
 Db 257 iqceplsapdlgimncshplasfstsactficsegteltgkrrkticessgimwnpspic 316
 OY 121 QKDKSFSMIKE 132
 Db 317 qkdkfsfmike 328

RESULT 21
 AAR98131
 ID AAR98131 standard; Protein; 374 AA.
 AC AAR98131;
 XX
 DT 01-NOV-1996 (first entry)
 XX

DE Human lymphocyte cell surface glycoprotein (HuLHR) variant.
 XX
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 XX

OS Homo sapiens.
 XX
 PN
 XX
 PD
 XX
 PF
 XX
 PR
 XX
 PR
 XX
 PR
 XX
 PR
 XX

Key Location/Qualifiers
 Region 1..37
 Domain /label= Signal region.
 Domain 38..157
 Domain /label= Lectin domain.
 Domain 162..195
 Domain /label= EGF domain.
 Binding-site 199..260
 Binding-site /label= Complement binding repeat 1.
 Binding-site 261..319
 Binding-site /label= Complement binding repeat 2.
 Domain 335..357
 Domain /label= Transmembrane domain.
 Domain 358..374
 Domain /label= Cytoplasmic domain.

[illegible]

```

Query Match
Best Local Similarity 100.0%; Score 733; DB 17; Length 374;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPPGNFSFSQCAFSCSEGNLTGIEETTCGPGNMSSPEPTCOV 60
DB 199 ceplapdlgmdcthpignfsfsqcafscegnltgiettcgpgnmsspeptcqv 258
QY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTFCSEGTLLICKKTTICSSGSIWNSPSPIC 120
DB 259 iqceplapdlgimncshplasfstsactfcsegtellgkktlccssgiwvnspspic 318
QY 121 QKLDKSFMSIKE 132
DB 319 qkldksfsmike 330

RESULT 23
AAR98122
ID AAR98122 standard; Protein; 371 AA.
AC AAR98122;
DT 01-NOV-1996 (first entry)
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
OS Homo sapiens.
XX
XX Key
XX Region 1..37 Location/Qualifiers
XX FT /label= Signal region.
XX FT /label= Lectin domain.
XX FT /label= EGF domain.
XX FT /label= EGF domain.
XX FT /label= Complement binding repeat 1.
XX FT /label= Complement binding repeat 2.
XX FT /label= Transmembrane domain.
XX FT /label= Transmembrane domain.
XX FT /label= Cytoplasmic domain.
XX PN US5514582-A.
XX PD 07-MAY-1996.
XX PF 23-FEB-1989; 89US-0315015.
XX PR 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
XX PR 21-JAN-1994; 94US-0185670.
XX PA (GENE ) GENENTECH INC.
XX PI Capon DJ, Lasky LA;
XX DR WPI; 1996-238773/24.
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to Ig constant region - useful for
XX diagnosis and treatment e.g. of inflammation
PS Disclosure; Page 19; 41pp; English.

```

```

XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically for the in
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection, inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an Ile288Val substitution.
SQ Sequence 371 AA;

```

```

Query Match
Best Local Similarity 99.9%; Score 732; DB 17; Length 371;
Matches 131; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPPGNFSFSQCAFSCSEGNLTGIEETTCGPGNMSSPEPTCOV 60
DB 196 ceplapdlgmdcthpignfsfsqcafscegnltgiettcgpgnmsspeptcqv 255
QY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTFCSEGTLLICKKTTICSSGSIWNSPSPIC 120
DB 256 iqceplapdlgimncshplasfstsactfcsegtellgkktlccssgiwvnspspic 315
QY 121 QKLDKSFMSIKE 132
DB 316 qkldksfsmike 327

RESULT 24
AAR98124
ID AAR98124 standard; Protein; 371 AA.
AC AAR98124;
DT 01-NOV-1996 (first entry)
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Key
XX Region 1..37 Location/Qualifiers
XX FT /label= Signal region.
XX FT /label= Lectin domain.
XX FT /label= EGF domain.
XX FT /label= EGF domain.
XX FT /label= Complement binding repeat 1.
XX FT /label= Complement binding repeat 2.
XX FT /label= Transmembrane domain.
XX FT /label= Transmembrane domain.
XX FT /label= Transmembrane domain.
XX PN US5514582-A.
XX PD 07-MAY-1996.
XX PF 23-FEB-1989; 89US-0315015.
XX PR 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
XX PR 21-JAN-1994; 94US-0185670.
XX PA (GENE ) GENENTECH INC.
XX PI Capon DJ, Lasky LA;
XX DR WPI; 1996-238773/24.
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to Ig constant region - useful for
XX diagnosis and treatment e.g. of inflammation
PS Disclosure; Page 19; 41pp; English.

```



```

FT      XX      /label= Cytoplasmic domain.
XX      PN      USS514582-A.
XX      PD      07-MAY-1996.
XX      PF      23-FEB-1989;      89US-0315015.
XX      PR      22-NOV-1989;      89US-0440625.
XX      PR      23-FEB-1989;      89US-0315015.
XX      PR      16-DEC-1991;      91US-0808122.
XX      PR      08-DEC-1992;      92US-0986931.
XX      PR      21-JAN-1994;      94US-0185670.
XX      PA      (GETH ) GENENTECH INC.
XX      PI      Capon DJ, Lasky LA:
XX      DR      WPI: 1996-238773/24.
XX      PT      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX      PT      binding site of a receptor fused to Ig constant region - useful for
XX      PS      diagnosis and treatment e.g. of inflammation
XX      PS      Disclosure: Page 19; 41pp; English.
XX      CC      A hybrid immunoglobulin chain comprising the ligand binding site of
XX      CC      a single transmembrane receptor without an active transmembrane
XX      CC      region, fused at its C-terminus with the N-terminus of an
XX      CC      immunoglobulin constant region. The receptor is not a member of the
XX      CC      immunoglobulin super family, nor a multiple subunit polypeptide
XX      CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
XX      CC      the adhesion/targeting of a ligand binding partner (LBP) with the
XX      CC      effector functions of immunoglobulin and can bind to and/or activate
XX      CC      more than one ligand. It can be used diagnostically for the in
XX      CC      vitro assay of LBP and their targets; or therapeutically to deliver
XX      CC      LBP such as toxins, enzymes, growth factors to particular cells.
XX      CC      Typical applications are as antiviral, neuromodulating and
XX      CC      immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX      CC      treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX      CC      The immunoglobulin component increases plasma half life and
XX      CC      facilitates purification while deletion of the transmembrane region
XX      CC      facilitates recovery, improves aqueous solubility and removes
XX      CC      potentially immunogenic epitopes. Variants of the human lymphocyte
XX      CC      cell surface glycoprotein described in AAR98106 are given in
XX      CC      AAR98109-R98135. This variant contains an Ile302Leu substitution.
XX      S0      Sequence      371 AA:

Query Match      99.7%; Score 731; DB 17; Length 371;
Best Local Similarity 99.2%; Pred. No. 2.2e-63;
Matches 131; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

OY      1      CEPLEAPETGTMDCHHPFEGNFSSQCAFSCEGTNLTGIEETGCFEGNMSSPEPTCOY 60
DB      196      cepleapelgtmdchhpfnfssqcafscegtnlgtieetogcfpfnwsspeptcy 255
OY      61      IOCEPLAPADLCIMNCSPHPLASFSFTSACTFTCSCEGTLLGKKKTKTCSSSGIWSNDSPIC 120
DB      256      lqceplaspdpdlmncshplasfstsactftelgllgkkktlccsssglwnspspic 315
OY      121      OKLDKSFMIKE 132
DB      316      qkldksfsmike 327

RESULT 25
AAR98119
ID      AAR98119 standard; Protein; 371 AA.
XX      AC      AAR98119;
XX

```

DT	01-NOV-1996	(first entry)
XX		
DE	Human lymphocyte cell surface glycoprotein (HLHR) variant.	
XX		
KM	Immunoglobulin: transmembrane receptor; adhesion; targeting;	
KM	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KM	immunomodulator; cell adhesion; graft rejection; inflammation;	
KM	metastasis.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= EGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1.
FT	Binding-site	258..316
FT		/label= Complement binding repeat 2.
FT	Domain	332..354
FT		/label= Transmembrane domain.
FT	Domain	355..371
FT		/label= Cytoplasmic domain.
XX		
PN	US5514582-A.	
XX		
FT	07-MAY-1996.	
PD		
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GETH) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI: 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
XX	diagnosis and treatment e.g. of inflammation	
PS		
XX	Disclosure: Page 19; 41pp; English.	
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain comprises	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAR98106 are given in	
CC	AAR98109-RB8135. This variant contains a Ser226Thr substitution.	
XX		
XX	Sequence 371 AA:	

secreted and encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation

Key	Location/qualifiers
FT	1..37
FT	/label= Signal region.
FT	38..154
FT	/label= Lectin domain.
FT	159..192
FT	/label= EGF domain.
FT	196..257
FT	/label= Complement binding repeat 1.
FT	258..316
FT	/label= Complement binding repeat 2.
FT	332..354
FT	Domain

FT	Domain	/label= Transmembrane domain. 355..371 /label= Cytoplasmic domain.
FT	US5514582-A.	
XX	07-MAY-1996.	
XX	23-FEB-1989;	89US-0315015.
XX	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX	(GETH) GENENTECH INC.	
XX	Capon DJ, Lasky LA;	
PI	WPI: 1996-238773/24.	
XX		
XX		
DR		
XX		
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of Inflammation	
XX		
XX		
PS	Disclosure; Page 19; 41pp; English.	
XX		
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region, fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAR98106 are given in	
CC	AAR98109-R98135. This variant contains a Thr282Ser substitution.	
XX		
XX	Sequence 371 AA;	
SQ		
Query Match	99.5%;	Score 729; DB 17; Length 371;
Best Local Similarity	99.2%;	Pred. No. 3,5e-63;
Matches 131; Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
QY	1	CEPFLAPELGTNDCTHPFGNFFSFSOCASFCSCEGNTLGIERTGCPFGNMSSPPTCOV 60
DB	196	ceplapeplgtndctchpfnfifsgqcafcscgcnlgiertcpgfnwaspsptcgv 255
QY	61	IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSGETELIGKRTICSSSGIWSNPSPIC 120
DB	256	iqceplsapdlgimncshplasfssactfcscgtelelgykkkticessgylwpspspic 315
QY	121	OKLQKSPFSNIKE 132
DB	316	qklqksfsmike 327

AC	AAK22802;	
DT	01-SEP-1992 (first entry)	
XX		
XX		
DE	Human lymphocyte homing receptor.	
XX		
KM	HuLHR; LHR; binding; endothelium; immunogens; graft; organ;	
XX	rejection; inflammation; rheumatoid arthritis; lymphoma metastasis.	
OS		
XX	Homo sapiens.	
PH		
FT	Key	Location/Qualifiers
FT	Region	20..32
FT		/note= "potential signal sequence"
FT	Modified-site	60..62
FT		/note= "N-glycosylation site"
FT	Modified-site	104..106
FT		/note= "N-glycosylation site"
FT	Modified-site	177..179
FT		/note= "N-glycosylation site"
FT	Modified-site	216..218
FT		/note= "N-glycosylation site"
FT	Modified-site	232..234
FT		/note= "N-glycosylation site"
FT	Modified-site	271..273
FT		/note= "N-glycosylation site"
FT	Modified-site	311..313
FT		/note= "N-glycosylation site"
FT	Region	335..357
FT		/note= "stop transfer or membrane anchoring domain"
PN		
XX	US5098833-A.	
XX		
XX	24-MAR-1992.	
PD		
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	23-FEB-1989;	89US-0315015.
XX		
PA	(RECC) UNIV OF CALIFORNIA.	
PI		
XX	Lasky LA, Rosen SD, Stachel SE, Singer MS, Yednock TA;	
XX		
DR	WPI; 1992-123385/15.	
XX		
DR	N-PSDB; AAQ23623.	
XX		
PT	New DNA encoding at least one domain of lymphocyte homing	
PT	receptor - useful for treating graft rejection, inflammation,	
PT	etc.	
PS		
XX	Disclosure: Fig 1; 32pp; English.	
XX		
CC	The protein sequence was deduced from the DNA sequence obtd. by	
CC	screening an oligo dt primed lambda gt10 cDNA library derived from	
CC	human peripheral blood lymphocyte mRNA obtd. from primary cells,	
CC	with a 2.2 kb EcoRI insert of the murine Mel 14 antigen cDNA clone.	
CC	The protein contains regions encoding a carbohydrate binding domain,	
CC	an EGF-like domain, a complement binding domain and a transmembrane	
CC	domain. The protein contains 26 cystein residues. Cells transfomed	
CC	by the hLHR DNA are used to produce LHR (which mediates binding of	
CC	lymphocytes to the endothelium of lymphoid tissue). LHR or its	
CC	variants are useful as reagents for assaying LHR or anti-LHR anti-	
CC	bodies, to purify the antibodies, as immunogens, and therapeutically	
CC	to compete with normal binding of lymphocytes (to prevent graft/organ	
CC	rejection, to treat inflammation (such as Rheumatoid arthritis or	
CC	other autoimmune diseases); for control of lymphocyte metastasis,	
CC	and to treat conditions associated with accumulation of lymphocytes),	
CC	derivs. and variants of LHR may be produced having modified	
CC	properties, e.g. increased activity, longer plasma half-life,	
CC	reduced side effects and better aq. solubility.	
CC	See also AAK22803.	
XX		
XX	Sequence 372 AA;	

Query Match 99.3%; Score 728; DB 13; Length 372;
 Best Local Similarity 99.2%; Pred. No. 4.4e-63;
 Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSFSSQCAFSCSEGNLTGIEETTCGPGNMSSPEPTQY 60
 |||||
 Db 197 ceplapdlgmdcthpngfsfsqcafscsegnltgieettcgpgnmspeptcqy 256
 |||||

QY 61 IQCEPLAPDLGIMNCSPPLASFSFTSACTGTCSEGTGLIGKKTICSSGIMWNSPSPIC 120
 |||||
 Db 257 iqceplapdlgimncspplasfsftsaacticsegtelgkkticessgimwnspspic 316
 |||||

QY 121 QKLDKSFMSIKE 132
 |||||
 Db 317 qkldksfsmike 328
 |||||

RESULT 29

AAR98118
 ID AAR98118 standard; Protein; 371 AA.

AC AAR98118;

DT 01-NOV-1996 (first entry)

DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.

OS Homo sapiens.

Key Location/Qualifiers
 FH Region 1..37
 FT /label= Signal region.
 FT Domain 38..154
 FT /label= Lectin domain.
 FT Domain 159..192
 FT /label= EGF domain.
 FT Binding-site 196..257
 FT /label= Complement binding repeat 1.
 FT Binding-site 258..316
 FT /label= Complement binding repeat 2.
 FT Domain 332..354
 FT /label= Transmembrane domain.
 FT Domain 355..371
 FT /label= Cytoplasmic domain.

PN USS514582-A.

PD 07-MAY-1996.

PF 23-FEB-1989; 89US-0315015.

PR 23-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

PA (GENE) GENENTECH INC.

PI Capon DJ, Lasky LA;

PT WPI; 1996-238773/24.
 Nucleic acid encoding hybrid immunoglobulin comprising the ligand
 binding site of a receptor fused to Ig constant region - useful for
 diagnosis and treatment e.g. of inflammation

XX
 PS Disclosure; Page 19; 41pp; English.

CC A hybrid immunoglobulin chain comprising the ligand binding site of
 CC a single transmembrane receptor without an active transmembrane
 CC region; fused at its C-terminus with the N-terminus of an
 CC immunoglobulin constant region. The receptor is not a member of the
 CC immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver
 CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC the immunoglobulin component increases metastasis of lymphoma etc.)
 CC facilitates purification while deletion of the transmembrane region
 CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-R98135. This variant contains a Phe214Ileu substitution.

CC Sequence 371 AA;

Query Match 99.2%; Score 727; DB 17; Length 371;
 Best Local Similarity 99.2%; Pred. No. 5.5e-63;

Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSFSSQCAFSCSEGNLTGIEETTCGPGNMSSPEPTQY 60
 |||||
 Db 196 ceplapdlgmdcthpngfsfsqcafscsegnltgieettcgpgnmspeptcqy 255
 |||||

QY 61 IQCEPLAPDLGIMNCSPPLASFSFTSACTGTCSEGTGLIGKKTICSSGIMWNSPSPIC 120
 |||||
 Db 256 iqceplapdlgimncspplasfsftsaacticsegtelgkkticessgimwnspspic 315
 |||||

QY 121 QKLDKSFMSIKE 132
 |||||
 Db 316 qkldksfsmike 327
 |||||

RESULT 30

AAR98120
 ID AAR98120 standard; Protein; 371 AA.

AC AAR98120;

DT 01-NOV-1996 (first entry)

DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.

OS Homo sapiens.

Key Location/Qualifiers
 FH Region 1..37
 FT /label= Signal region.
 FT Domain 38..154
 FT /label= Lectin domain.
 FT Domain 159..192
 FT /label= EGF domain.
 FT Binding-site 196..257
 FT /label= Complement binding repeat 1.
 FT Binding-site 258..316
 FT /label= Complement binding repeat 2.
 FT Domain 332..354

```

FT      Domain                               /label= Transmembrane domain.
FT      355..371
FT      /label= Cytoplasmic domain.
XX
XX
XX      US5514582-A.
XX
XX      07-MAY-1996.
XX
XX      23-FEB-1989;      89US-0315015.
XX
XX      22-NOV-1989;      89US-0440625.
XX      23-FEB-1989;      89US-0315015.
XX      16-DEC-1991;      91US-0808122.
XX      08-DEC-1992;      92US-0986931.
XX      21-JAN-1994;      94US-0185670.
XX
XX      (GETH ) GENENTECH INC.
XX
XX      Capon DJ, Lasky LA;
XX
XX      WPI: 1996-238773/24.
XX
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX      binding site of a receptor fused to Ig constant region - useful for
XX      diagnosis and treatment e.g. of inflammation
XX
XX      Disclosure: Page 19; 41pp; English.
XX
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
XX      a single transmembrane receptor without an active transmembrane
XX      region; fused at its C-terminus with the N-terminus of an
XX      immunoglobulin constant region. The receptor is not a member of the
XX      immunoglobulin super family, nor a multiple subunit polypeptide
XX      encoded by discrete genes. The hybrid immunoglobulin chain combines
XX      the adhesion/targeting of a ligand binding partner (LBP) with the
XX      effector functions of immunoglobulin and can bind to and/or activate
XX      more than one ligand. It can be used diagnostically for the in
XX      vitro assay of LBP and their targets; or therapeutically to deliver
XX      LBP such as toxins, enzymes, growth factors to particular cells.
XX      Typical applications are as antiviral, neuromodulating and
XX      immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX      treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX      The immunoglobulin component increases plasma half life and
XX      facilitates purification while deletion of the transmembrane region
XX      facilitates recovery, improves aqueous solubility and removes
XX      potentially immunogenic epitopes. Variants of the human lymphocyte
XX      cell surface glycoprotein described in AAR98106 are given in
XX      AAR98109-R98135. This variant contains a Phe244Met substitution.
XX
XX      Sequence      371 AA:
SQ

```

```

AC      AAR98123;
XX
XX      01-NOV-1996 (first entry)
XX
XX      Human lymphocyte cell surface glycoprotein (HLHR) variant.
XX
XX      Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX      diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX      immunomodulator; cell adhesion; graft rejection; inflammation;
XX      metastasis.
XX
XX      Homo sapiens.
XX
XX      OS
XX
XX      Key
XX      Region
XX      Domain
XX      Domain
XX      Domain
XX      Binding-site
XX      Binding-site
XX      Binding-site
XX      Binding-site
XX      Domain
XX      Domain
XX      Domain
XX      Domain
XX      Domain
XX      US5514582-A.
XX
XX      07-MAY-1996.
XX
XX      23-FEB-1989;      89US-0315015.
XX
XX      22-NOV-1989;      89US-0440625.
XX      23-FEB-1989;      89US-0315015.
XX      16-DEC-1991;      91US-0808122.
XX      08-DEC-1992;      92US-0986931.
XX      21-JAN-1994;      94US-0185670.
XX
XX      (GETH ) GENENTECH INC.
XX
XX      Capon DJ, Lasky LA;
XX
XX      WPI: 1996-238773/24.
XX
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX      binding site of a receptor fused to Ig constant region - useful for
XX      diagnosis and treatment e.g. of inflammation
XX
XX      Disclosure: Page 19; 41pp; English.
XX
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
XX      a single transmembrane receptor without an active transmembrane
XX      region; fused at its C-terminus with the N-terminus of an
XX      immunoglobulin constant region. The receptor is not a member of the
XX      immunoglobulin super family, nor a multiple subunit polypeptide
XX      encoded by discrete genes. The hybrid immunoglobulin chain combines
XX      the adhesion/targeting of a ligand binding partner (LBP) with the
XX      effector functions of immunoglobulin and can bind to and/or activate
XX      more than one ligand. It can be used diagnostically for the in
XX      vitro assay of LBP and their targets; or therapeutically to deliver
XX      LBP such as toxins, enzymes, growth factors to particular cells.
XX      Typical applications are as antiviral, neuromodulating and
XX      immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX      treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX      The immunoglobulin component increases plasma half life and
XX      facilitates purification while deletion of the transmembrane region
XX      facilitates recovery, improves aqueous solubility and removes
XX      potentially immunogenic epitopes. Variants of the human lymphocyte
XX      cell surface glycoprotein described in AAR98106 are given in
XX      AAR98109-R98135. This variant contains a Lys298-Lys299; Arg-Arg

```


AAR98135
 ID AAR98135 standard; Protein; 374 AA.
 AC AAR98135;
 XX
 DT 01-NOV-1996 (first entry)
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
 XX
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
 KW immunomodulator; cell adhesion; graft rejection; inflammation;
 KW metastasis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..37
 FT /label= Signal region.
 FT Domain 38..154
 FT /label= Lectin domain.
 FT Domain 159..192
 FT /label= EGF domain.
 FT Binding-site 196..257
 FT /label= Complement binding repeat 1.
 FT Binding-site 258..319
 FT /label= Complement binding repeat 2.
 FT 335..357
 FT /label= Transmembrane domain.
 FT Domain 358..374
 FT /label= Cytoplasmic domain.
 XX
 PN US5514582-A.
 XX
 PD 07-MAY-1996.
 XX
 PF 23-FEB-1989; 89US-0315015.
 XX
 PR 22-NOV-1989; 89US-0440625.
 PR 23-FEB-1989; 89US-0315015.
 PR 16-DEC-1991; 91US-0808122.
 PR 08-DEC-1992; 92US-0986931.
 PR 21-JAN-1994; 94US-0185670.
 XX
 PA (GENE) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 DR WPI; 1996-238773/24.
 XX
 PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
 PT binding site of a receptor fused to Ig constant region - useful for
 PT diagnosis and treatment e.g. of inflammation
 XX
 PS Disclosure; Page 19; 41pp; English.
 XX
 CC A hybrid immunoglobulin chain comprising the ligand binding site of
 CC a single transmembrane receptor without an active transmembrane
 CC region; fused at its C-terminus with the N-terminus of an
 CC immunoglobulin constant region. The receptor is not a member of the
 CC immunoglobulin super family, nor a multiple subunit polypeptide
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines
 CC the adhesion/targeting of a ligand binding partner (LBP) with the
 CC effector functions of immunoglobulin and can bind to and/or activate
 CC more than one ligand. It can be used diagnostically for the in
 CC vitro assay of LBP and their targets; or therapeutically to deliver
 CC LBP such as toxins, enzymes, growth factors to particular cells.
 CC Typical applications are as antiviral, neuromodulating and
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
 CC The immunoglobulin component increases plasma half life and
 CC facilitates purification while deletion of the transmembrane region
 CC facilitates recovery, improves aqueous solubility and removes

CC potentially immunogenic epitopes. Variants of the human lymphocyte
 CC cell surface glycoprotein described in AAR98106 are given in
 CC AAR98109-R98135. This variant contains a 292-Tyr-Tyr-Tyr insertion.
 XX
 SQ Sequence 374 AA;
 Query Match 98.4%; Score 721.5; DB 17; Length 374;
 Best Local Similarity 97.8%; Pred. No. 1,9e-62;
 Matches 132; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
 QY 1 CEPLEAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGFGMWSPEPTCY 60
 DB 196 cepleapdlgmcdcthpfnfssqcafscegnltgiettcgpfgnwsspeptcy 255
 QY 61 IOCEPLAPDLGIMNGCHPLASFSTACPTFCE--GTELGKKTICSSGIMWNP 117
 DB 256 iqceplapdlgimncshplastfstaacticeyygtellgkkticessgimw 315
 QY 118 PICQKLDKSFSSMIKE 132
 DB 316 picqkldksfsmike 330
 RESULT 36
 AAR22551
 ID AAR22551 standard; Protein; 363 AA.
 XX
 AC AAR22551;
 XX
 DT 21-MAY-1992 (first entry)
 XX
 DE T lymphocyte-specific Leu8 Antigen minor form.
 XX
 KW Rapid immunoselection cloning technique; cell surface antigen;
 KW homing receptor; antigen-presenting cells.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 113..115
 FT /label= "N-linked glycosylation"
 FT /note= "putative"
 FT Modified-site 157..159
 FT /label= "N-linked glycosylation"
 FT /note= "putative"
 FT Modified-site 230..232
 FT /label= "N-linked glycosylation"
 FT /note= "putative"
 FT Modified-site 269..271
 FT /label= "N-linked glycosylation"
 FT /note= "putative"
 FT Modified-site 285..287
 FT /label= "N-linked glycosylation"
 FT /note= "putative"
 FT Modified-site 299..301
 FT /label= "N-linked glycosylation"
 FT /note= "putative"
 FT Modified-site 324..326
 FT /label= "N-linked glycosylation"
 FT /note= "putative"
 W09201049-A.
 PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90WO-US04986.
 XX
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GENE) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Amlot M;

XX WPI: 1992-056864/07.
DR N-PSDB: AAQ22500.
XX
PT New CD53 cell surface antigen and DNA encoding it - for
PT Immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
PS Example 14; Page 106; 160pp; English.
XX
CC Two cDNA clones encoding Leu8 determinants were isolated from a
CC human T cell library using the rapid immunoselection cloning method
CC (see e.g. AAQ21164 for description of method). This protein sequence
CC was deduced from the shorter insert. The weakly hydrophobic
CC C-terminal domain is characteristic of surface proteins that are
CC attached to the cell membrane by covalent linkage to a
CC phosphatidylinositol-substituted glycan.
CC See AAQ21184 for the larger insert and AAR20815 for the major form of
CC the Leu8 antigen that it encodes.
XX
SQ Sequence 363 AA:

Query Match 98.1%; Score 719; DB 13; Length 363;
Best Local Similarity 98.5%; Pred. No. 3.2e-62;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGTGIEETTCGPRGNSSPEPTQY 60
DB 210 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGTGIEETTCGPRGNSSPEPTQY 269
QY 61 IQCEPLAPDLGIMNCSHPLASFSTACTFICSEGTLELIGKKKTICSSGIMNSPPTIC 120
DB 270 IQCEPLAPDLGIMNCSHPLASFSTACTFICSEGTLELIGKKKTICSSGIMNSPPTIC 329
QY 121 QKDKSFSMIKE 132
DB 330 QKDKSFSMIKE 341

RESULT 37
AAR91443
ID AAR91443 standard; Protein: 363 AA.
XX
AC AAR91443;
XX
DT 31-OCT-1996 (first entry)
XX
DE Human Leu8 antigen.
XX
KW Cell surface antigen; cloning; immunoselection; immunotherapy;
KW therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;
KW antiinflammatory.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 113..115
FT Modified-site /label= glycosylation_site
FT Modified-site 156..158
FT Modified-site /label= glycosylation_site
FT Modified-site 229..231
FT Modified-site /label= glycosylation_site
FT Modified-site 269..271
FT Modified-site /label= glycosylation_site
FT Modified-site 285..287
FT Modified-site /label= glycosylation_site
FT Modified-site 299..301
FT Modified-site /label= glycosylation_site
FT Modified-site 324..326
FT Modified-site /label= glycosylation_site
XX
XX US506126-A.

PD 09-APR-1996.
XX
XX 25-FEB-1988; 88US-0160416.
XX
PR 01-DEC-1992; 92US-0983647.
PR 25-FEB-1988; 88US-0160416.
PR 13-JUL-1989; 89US-0379076.
PR 13-JUL-1990; 90US-0553759.
PR 18-OCT-1993; 93US-0139273.
XX
XX (GENO) GEN HOSPITAL CORP.
XX
PI Aruffo A, Seed B;
XX
XX WPI: 1996-200279/20.
DR N-PSDB: AAT14723.
XX
PT Cloning of cDNA encoding cell surface antigen - useful for isolation
PT of diagnostic and therapeutic proteins
XX
PS Example 14; Column 71-74; 79pp; English.
XX
CC 2 cDNA clones encoding Leu8 determinants (AAR91442 and AAR91443) were
CC isolated from a human T-cell library using a novel immunoselection
CC cloning method. The longer insert (AAR14723) contained 2,350
CC residues, while the shorter lacked 436 internal residues. A major
CC transcript of 2.4 kb was present in peripheral blood mononuclear
CC cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,
CC and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukaemic
CC T-cell lines. Leu8 antigens can be obtd. for diagnostic and
CC therapeutic use. The presence or absence of Leu8 on CD4+ T-cells
CC identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.
CC Soluble forms of Leu8 can act as antiinflammatory agents by reducing
CC lymphocyte migration.
XX
SQ Sequence 363 AA:

Query Match 98.1%; Score 719; DB 17; Length 363;
Best Local Similarity 98.5%; Pred. No. 3.2e-62;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGTGIEETTCGPRGNSSPEPTQY 60
DB 210 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGTGIEETTCGPRGNSSPEPTQY 269
QY 61 IQCEPLAPDLGIMNCSHPLASFSTACTFICSEGTLELIGKKKTICSSGIMNSPPTIC 120
DB 270 IQCEPLAPDLGIMNCSHPLASFSTACTFICSEGTLELIGKKKTICSSGIMNSPPTIC 329
QY 121 QKDKSFSMIKE 132
DB 330 QKDKSFSMIKE 341

RESULT 38
AAY96184
ID AAY96184 standard; Protein: 363 AA.
XX
AC AAY96184;
XX
DT 19-DEC-2000 (first entry)
XX
DE Human T-cell specific Leu8 antigen.
XX
KW Leu8; cell surface antigen; human; immunoselection; panning;
KW immunodiagnosis; diagnosis; immunotherapy; gene therapy;
KW immune disorder; infection; asthma; immune-complex disease;
KW amyloidosis; multiple sclerosis; inflammation; antiinflammatory.
XX
XX Homo sapiens.
XX
XX US611093-A.


```

QY 1 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGIRETTGCGPRGNMSSPEPTQV 60
DB 210 cepleaprlgmdcthsignfsfsqcafscsegtnlgiectcpgfnwsspeptqv 269
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTLELIGKKTKTICSSGIMSNPSPIC 120
DB 270 lqceplaspdlgimncshplasfstsactficsgtelgkktkicssglwnspspic 329
QY 121 QKLDKSFMSIKE 132
DB 330 qkldksfsmike 341

RESULT 40
AAR91442
ID AAR91442 standard; Protein: 385 AA.
XX
AC AAR91442;
XX
DT 31-OCT-1996 (first entry)
XX
DE Human Leu8 antigen.
XX
KW Cell surface antigen; cloning: immunoselection; immunotherapy;
KW therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;
XX
XX antinflammatory.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT Modified-site 113..115
FT Modified-site /label= Glycosylation_site
FT Modified-site 136..158
FT Modified-site /label= Glycosylation_site
FT Modified-site 229..231
FT Modified-site /label= Glycosylation_site
FT Modified-site 269..271
FT Modified-site /label= Glycosylation_site
FT Modified-site 285..287
FT Modified-site /label= Glycosylation_site
FT Modified-site 299..301
FT Modified-site /label= Glycosylation_site
FT Modified-site 324..326
FT Modified-site /label= Glycosylation_site
FT Domain 347..368
FT /label= Transmembrane_domain
XX
XX US5506126-A.
XX
XX 09-APR-1996.
XX
XX 25-FEB-1988; 88US-0160416.
XX
XX 01-DEC-1992; 92US-0983647.
XX 25-FEB-1988; 88US-0160416.
XX 13-JUL-1989; 89US-0379076.
XX 13-JUL-1990; 90US-0553759.
XX 18-OCT-1993; 93US-0139273.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Arufo A, Seed B;
XX
XX WPI: 1996-200279/20.
XX N-PSDB; AAT14723.
XX
XX Cloning of cDNA encoding cell surface antigen - useful for isolation
XX of diagnostic and therapeutic proteins
XX
XX Example 14; Column 71-74; 79pp; English.
XX
XX 2 cDNA clones encoding Leu8 determinants (AAR91442 and AAR91443) were

```

```

CC isolated from a human T-cell library using a novel immunoselection
CC cloning method. The longer insert (AAT14723) contained 2,350
CC residues, while the shorter lacked 436 internal residues. A major
CC transcript of 2.4 kb was present in peripheral blood mononuclear
CC cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,
CC and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukemia
CC T-cell lines. Leu8 antigens can be obtd. for diagnostic and
CC therapeutic use. The presence or absence of Leu8 on CD4+ T-cells
CC identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.
CC Soluble forms of Leu8 can act as antinflammatory agents by reducing
CC lymphocyte migration.
XX
XX Sequence 385 AA;
SO

```

```

Query Match 98.1%; Score 719; DB 17; Length 385;
Best Local Similarity 98.5%; Pred. No. 3,56-62;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGIRETTGCGPRGNMSSPEPTQV 60
DB 210 cepleaprlgmdcthsignfsfsqcafscsegtnlgiectcpgfnwsspeptqv 269
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFICSEGTLELIGKKTKTICSSGIMSNPSPIC 120
DB 270 lqceplaspdlgimncshplasfstsactficsgtelgkktkicssglwnspspic 329
QY 121 QKLDKSFMSIKE 132
DB 330 qkldksfsmike 341

```

```

Search completed: September 7, 2002, 10:14:09
Job time: 267 sec

```

This Page Blank (uspto)

This Page Blank (uspto)